

Ann. Henri Poincaré 18 (2017), 3049–3096  
© 2017 The Author(s). This article is published  
with open access at Springerlink.com  
1424-0637/17/093049-48  
published online March 13, 2017  
DOI 10.1007/s00023-017-0566-1

Annales Henri Poincaré



# Causality for Nonlocal Phenomena

Michał Eckstein and Tomasz Miller

**Abstract.** Drawing from the theory of optimal transport we propose a rigorous notion of a causal relation for Borel probability measures on a given spacetime. To prepare the ground, we explore the borderland between Lorentzian geometry, topology and measure theory. We provide various characterisations of the proposed causal relation, which turn out to be equivalent if the underlying spacetime has a sufficiently robust causal structure. We also present the notion of the ‘Lorentz–Wasserstein distance’ and study its basic properties. Finally, we outline the possible applications of the developed formalism in both classical and quantum physics.

## 1. Introduction

The notion of a space, understood as a set of points, provides an indispensable framework for every physical theory. But, regardless of the physical system that is being modelled, the space itself is not directly observable. Indeed, any measuring apparatus can provide information about the localisation only up to a finite resolution. In the relativistic context, it means that the *event* is an idealised concept, which is not accessible to any observer.

Apart from the ‘practical’ obstructions for measuring position, there exist also fundamental ones because of the quantum effects manifest at small scales. Although nonrelativistic quantum mechanics does not impose any *a priori* restrictions on the accuracy of the position measurement, in quantum field theory a suitable ‘position operator’ is always nonlocal (see for instance [15, 24, 56]). Moreover, an attempt to perform a very accurate measurement of localisation in spacetime would require the use of signals of very short wavelength, resulting in an extreme concentration of energy. The latter would eventually lead to black hole formation, and the desired information would become trapped [19, 20].

It is generally believed that any physical theory should be causal, i.e. that no information can be transmitted with the speed exceeding the velocity of light. On a flat spacetime, one can develop a theory of tachyons consistent with special relativity, but it would have undesirable physical properties, for instance the vacuum state would not be Lorentz-invariant [23]. In general relativity, although Einstein's equations admit spacetime solutions with closed causal curves, these lead to paradoxes and are usually discarded as unphysical [34]. The status of causality in quantum theory was controversial from its dawn because of the inherent nonlocality of quantum wave functions [22]. However, careful studies in quantum information theory have proven that quantum nonlocality on its own cannot be utilised for a superluminal transfer of information [49]. In fact, the request of no faster-than-light signalling is often used as a guiding principle to restrain the admissible quantum theories [6, 32] and their possible extensions [47]. In quantum field theory, where nonlocality is even more prevailing, the postulate of causality is promoted to one of the axioms [33, 54]. Some of the approaches to quantum gravity suggest the breakdown of Einstein's causality at the level of Planck scale [4], but no evidence of such a phenomenon has been found so far [1].

In relativity theory it is straightforward to implement the postulate of causality as the Lorentzian metric induces a precise notion of causal curves between the point events. However, when a physical object requires a nonlocal description, the notion of a causal relation becomes hazy. This is particularly pertinent in quantum mechanics, as for instance the interference fringes resulting from quantum superposition can travel superluminally, but cannot be utilised to send information [9]. In fact, different concepts of causality in quantum theory can lead to controversial results—compare for instance [36] and [16] or [2] and [37, 62]. Hence, the issue of causality for nonlocal phenomena is still a timely subject of study [3, 5, 9, 17, 32, 38, 47, 49, 56, 60, 62].

The aim of this paper is to provide a rigorous notion of a causal relation between probability measures on a given spacetime. These can be utilised to model various nonlocal phenomena ranging from classical dust densities in cosmology, through energy or charge distributions in relativistic continuum mechanics, to quantum probability densities arising from the 'modulus square principle' in the wave packet formalism. Moreover, one can make use of probability measures to take into account experimental errors, as the measurement of any physical object's spacetime localisation would effectively be vitiated by an error resulting from the apparatus' imperfection.

We shall work in a generally covariant framework and keep the discourse in the spirit of mathematical relativity. We allow the probability measures to spread also in the timelike direction to facilitate applications that reach beyond the standard evolutionary approach.

The paper is organised as follows: in Sect. 2 we recall some basic notions in topology, measure theory and causality, to make the paper self-contained and accessible to a broad range of researchers. Section 3 contains the first

result on the verge of causality and measure theory—we demonstrate the  $\sigma$ -compactness of the set  $J^+ \subseteq \mathcal{M}^2$  in any spacetime. It establishes the foundations for the formalism developed in the next sections, but it also might be of independent interest for a mathematical-relativity-oriented reader.

The main concepts and results of the paper are aggregated in Sect. 4. We start off with a ‘dual’ definition of the causal relation, based on the notion of *causal functions* [44, Definition 2.3], proposed in [25] in a much wider context of noncommutative geometry. In several steps we show that it encapsulates an intuitive notion of causality for nonlocal objects:

Each infinitesimal part of the probability distribution should travel along a future-directed causal curve.

At each step we keep the causality conditions imposed on the underlying spacetime as low as possible. At the same time we provide several characterisations of causality for probability measures, which illustrate the concept and provide tools for concrete computations.

We are eventually led to the theory of optimal transport adapted to the relativistic setting. The latter is a new and fast-developing area of research [10, 13, 55],<sup>1</sup> which has found successful applications in the early universe reconstruction problem [14, 28–30].

Motivated by the main result, we put forward in Sect. 4.2 a definition of a causal relation between probability measures, valid on any spacetime, and study its properties. In particular, we demonstrate that the proposed relation is a partial order in the space of Borel probability measures on a given spacetime  $\mathcal{M}$ , even with a relatively poor causal structure.

Finally, we propose in Sect. 5 a notion of the ‘Lorentz–Wasserstein’ distance in the space of measures, enforcing the established bridge between mathematical physics and the theory of optimal transport.

We conclude, in Sect. 6, with an outlook into the possible future developments and applications. In particular, we briefly discuss the potential use of the presented results in the study of causality in quantum theory. We also address the interrelation of probability measures with states on  $C^*$ -algebras. In this way we provide a connection with the notion of ‘causality in the space of states’ proposed originally in [25] in the framework of noncommutative geometry.

## 2. Preliminaries

Throughout the paper we denote  $\mathbb{N} := \{1, 2, \dots\}$  and  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ .

The spaces of continuous, continuous and bounded, continuous and compactly supported real-valued functions on a topological space  $\mathcal{M}$  will be, respectively, denoted by  $C(\mathcal{M})$ ,  $C_b(\mathcal{M})$ ,  $C_c(\mathcal{M})$ . Analogous spaces of smooth functions will be, respectively, denoted by  $C^\infty(\mathcal{M})$ ,  $C_b^\infty(\mathcal{M})$ ,  $C_c^\infty(\mathcal{M})$ .

---

<sup>1</sup> The preprint [55] appeared on arXiv shortly after the first version of our paper (arXiv:1510.06386). Its motivation was somewhat different than ours, aiming at the optimal transport problems on globally hyperbolic Lorentz–Finsler manifolds. It independently led the author to a similar concept of a causal relation between probability measures.

## 2.1. Topology

Let  $\mathcal{M}$  denote a topological space and let  $\mathcal{X} \subseteq \mathcal{M}$ . The closure, interior, boundary and complement of  $\mathcal{X}$  will be, respectively, denoted by  $\overline{\mathcal{X}}$ ,  $\text{int } \mathcal{X}$ ,  $\partial \mathcal{X}$  and  $\mathcal{X}^c$ .

An *open cover* of  $\mathcal{X} \subseteq \mathcal{M}$  is a family  $\{U_\alpha\}_{\alpha \in A}$  of open subsets of  $\mathcal{M}$  such that  $\bigcup_{\alpha \in A} U_\alpha \supseteq \mathcal{X}$ .  $\mathcal{M}$  is called *Lindelöf* iff each of its open covers has a countable subcover.

A subset  $\mathcal{X} \subseteq \mathcal{M}$  is called *compact* iff each of its open covers has a finite subcover. It is called *sequentially compact* iff every sequence in  $\mathcal{X}$  has a subsequence convergent in  $\mathcal{X}$ . It is called *precompact* (or *relatively compact*) iff its closure is compact. Finally, it is called  $\sigma$ -*compact* iff it is a countable union of compact subsets. In particular,  $\mathcal{M}$  is  $\sigma$ -compact if and only if it admits an *exhaustion by compact sets*, that is a sequence  $(K_n)_{n \in \mathbb{N}}$  of compact sets such that  $K_n \subseteq K_{n+1}$  and  $\bigcup_{n=1}^\infty K_n = \mathcal{M}$ .

If  $\mathcal{M}$  is Hausdorff, then each of its compact subsets is closed. If  $\mathcal{M}$  is *second-countable*, that is if  $\mathcal{M}$  has a countable base, then the notions of compactness and sequential compactness coincide.

A Hausdorff<sup>2</sup> space  $\mathcal{M}$  is called *locally compact* iff each of its points has a precompact neighbourhood.

$\mathcal{M}$  is called *separable* iff there exists a countable subset  $\{a_n\}_{n \in \mathbb{N}} \subseteq \mathcal{M}$  dense in  $\mathcal{M}$ . Every open subspace of a separable space is itself separable.  $\mathcal{M}$  is called (*completely*) *metrisable* iff there exists a (complete) metric  $\rho : \mathcal{M}^2 \rightarrow \mathbb{R}_{\geq 0}$  inducing its topology. Fixing a metric allows one to talk about *balls*. By  $B(x, \varepsilon) := \{y \in \mathcal{M} \mid \rho(x, y) < \varepsilon\}$  we denote an *open ball* centred at  $x \in \mathcal{M}$  of radius  $\varepsilon > 0$ . By  $\overline{B}(x, \varepsilon)$  we denote its closure. Finally,  $\mathcal{M}$  is called *Polish* iff it is separable and completely metrisable.

In the following, we are going to work with *spacetimes* (see Sect. 2.3), which are examples of second-countable locally compact Hausdorff (LCH) spaces. Every such space is

- Lindelöf [61, Theorem 16.9];
- Polish, because [61, Theorems 19.3 & 23.1] imply that every second-countable LCH is metrisable, and by [52, Corollary 2.3.32] this means that every second-countable LCH is Polish;
- $\sigma$ -compact, because by taking a countable set  $\{a_n\}_{n \in \mathbb{N}}$  dense in  $\mathcal{M}$  (which always exists by the separability of  $\mathcal{M}$ ), and denoting by  $U_n$  a precompact neighbourhood of  $a_n$  (which always exists by the local compactness of  $\mathcal{M}$ ), one has that  $\mathcal{M} = \bigcup_{n=1}^\infty \overline{U}_n$ .

Moreover, every open subspace of a second-countable LCH space is itself a second-countable LCH space [61, Theorem 18.4].

LCH spaces satisfy the somewhat modified version of Urysohn's lemma [51, 2.12].

<sup>2</sup> There exist many definitions of local compactness, which are all equivalent in Hausdorff spaces.

**Theorem 1.** (Urysohn's lemma, LCH version) *Let  $\mathcal{M}$  be a LCH space and let  $\mathcal{K} \subseteq U \subseteq \mathcal{M}$ , where  $\mathcal{K}$  is compact and  $U$  is open. Then, there exists  $f \in C_c(\mathcal{M})$  such that  $f|_{\mathcal{K}} \equiv 1$ ,  $0 \leq f \leq 1$  and  $\text{supp } f \subseteq U$ .*

To conclude this section, recall that a topological space is *connected* iff it is not a union of two disjoint nonempty open sets. Furthermore, it is *locally connected* iff it has a base of connected sets. In general, these two properties are independent from each other. Spacetimes, however, are both connected and locally connected, what has the following interesting consequence.

**Lemma 1.** *Let  $\mathcal{M}$  be a connected, locally connected, second-countable LCH space. Then  $\mathcal{M}$  admits an exhaustion by connected compact sets.*

*Proof.* By the separability of  $\mathcal{M}$  there exists a countable dense subset  $\{a_n\}_{n \in \mathbb{N}} \subseteq \mathcal{M}$ . Let  $U_n$  denote a precompact neighbourhood of  $a_n$ , the existence of which is guaranteed by the local compactness of  $\mathcal{M}$ . By the local connectedness of  $\mathcal{M}$ , for every  $n \in \mathbb{N}$  there exists an open connected set  $V_n$  such that  $a_n \in V_n \subseteq U_n$ . Note that every  $V_n$  is precompact, because  $\overline{V}_n$ , being a closed subset of a compact set  $\overline{U}_n$ , is itself compact.

Thus, the family  $\{V_n\}_{n \in \mathbb{N}}$  is a countable cover of  $\mathcal{M}$  by open, precompact, connected sets. Using this cover, one can construct an exhaustion of  $\mathcal{M}$  with the desired properties.

Concretely, define an increasing sequence  $(C_n)_{n \in \mathbb{N}}$  of subsets of  $\mathcal{M}$  recursively as

$$C_1 := V_1, \quad C_{n+1} := C_n \cup V_{i_n}, \quad (n \in \mathbb{N}),$$

where  $V_{i_n}$  is chosen in such a way that  $i_n \notin \{i_1, \dots, i_{n-1}\}$  and  $V_{i_n} \cap C_n \neq \emptyset$ . By the connectedness of  $\mathcal{M}$ , such  $V_{i_n}$  can always be found.

Indeed, suppose that for certain  $n \in \mathbb{N}$  all  $V_i$ s with  $i \notin \{i_1, \dots, i_{n-1}\}$  are disjoint with  $C_n$ . But  $C_n$  is a union of open sets, and since it is disjoint with  $\bigcup_{i \notin \{i_1, \dots, i_{n-1}\}} V_i$ , we would just write  $\mathcal{M}$  as a union of two disjoint nonempty open sets, contradicting the connectedness of  $\mathcal{M}$ .

Obviously, every  $C_n$  is connected. Since the closure of a connected set is itself connected, the sequence  $(\overline{C}_n)_{n \in \mathbb{N}}$  is an exhaustion of  $\mathcal{M}$  by connected compact sets.  $\square$

## 2.2. Measure Theory

Let  $\mathcal{M}$  be a topological space. The  $\sigma$ -algebra of Borel sets  $\mathcal{B}(\mathcal{M})$  is the smallest family of subsets of  $\mathcal{M}$  containing the open sets, which is closed under complements and countable unions (and hence also under countable intersections). If  $\mathcal{M}$  is a Hausdorff space, then, in particular, its  $\sigma$ -compact subsets are Borel.

A function  $f : \mathcal{M}_1 \rightarrow \mathcal{M}_2$  between topological spaces is called *Borel* iff  $f^{-1}(V) \in \mathcal{B}(\mathcal{M}_1)$  for any  $V \in \mathcal{B}(\mathcal{M}_2)$ . Every continuous (or even semi-continuous) real-valued function is Borel.

A *Borel probability measure* on  $\mathcal{M}$  is a function  $\mu : \mathcal{B}(\mathcal{M}) \rightarrow \mathbb{R}_{\geq 0}$  satisfying  $\mu(\mathcal{M}) = 1$  and  $\mu(\bigcup_{n=1}^{\infty} \mathcal{X}_n) = \sum_{n=1}^{\infty} \mu(\mathcal{X}_n)$  for any  $\{\mathcal{X}_n\}_{n \in \mathbb{N}} \subseteq \mathcal{B}(\mathcal{M})$  such that  $\mathcal{X}_n \cap \mathcal{X}_m = \emptyset$  for  $n \neq m$ . A Borel set, the measure  $\mu$  of which is zero,

is called  $\mu$ -null. The pair  $(\mathcal{M}, \mu)$  is called a *probability space*. The set of Borel probability measures on  $\mathcal{M}$  will be denoted by  $\mathfrak{P}(\mathcal{M})$ .

Every  $\mu \in \mathfrak{P}(\mathcal{M})$  has the following properties [51, Theorem 1.19]:

- $\mu(\emptyset) = 0$ ;
- $\mu$  is *monotone*, i.e.  $\forall \mathcal{X}_1, \mathcal{X}_2 \in \mathcal{B}(\mathcal{M}) \quad \mathcal{X}_1 \subseteq \mathcal{X}_2 \Rightarrow \mu(\mathcal{X}_1) \leq \mu(\mathcal{X}_2)$ ;
- $\mu$  is *countably subadditive*, i.e.

$$\forall \{\mathcal{X}_n\}_{n \in \mathbb{N}} \subseteq \mathcal{B}(\mathcal{M}) \quad \mu \left( \bigcup_{n=1}^{\infty} \mathcal{X}_n \right) \leq \sum_{n=1}^{\infty} \mu(\mathcal{X}_n);$$

- for any sequence  $(\mathcal{X}_n)_{n \in \mathbb{N}} \subseteq \mathcal{B}(\mathcal{M})$  which is *increasing*, i.e.  $\mathcal{X}_n \subseteq \mathcal{X}_{n+1}$ , it is true that

$$\mu \left( \bigcup_{n=1}^{\infty} \mathcal{X}_n \right) = \lim_{n \rightarrow +\infty} \mu(\mathcal{X}_n); \quad (1)$$

- for any sequence  $(\mathcal{X}_n)_{n \in \mathbb{N}} \subseteq \mathcal{B}(\mathcal{M})$  which is *decreasing*, i.e.  $\mathcal{X}_n \supseteq \mathcal{X}_{n+1}$ , it is true that

$$\mu \left( \bigcap_{n=1}^{\infty} \mathcal{X}_n \right) = \lim_{n \rightarrow +\infty} \mu(\mathcal{X}_n). \quad (2)$$

Furthermore, if  $\mathcal{M}$  is metrisable, then every  $\mu \in \mathfrak{P}(\mathcal{M})$  is *regular* [52, Lemma 3.4.14], i.e.

$$\begin{aligned} \forall \mathcal{X} \in \mathcal{B}(\mathcal{M}) \quad \mu(\mathcal{X}) &= \sup \{ \mu(F) \mid F \subseteq \mathcal{X}, F \text{ closed} \} \\ &= \inf \{ \mu(U) \mid U \supseteq \mathcal{X}, U \text{ open} \}. \end{aligned} \quad (3)$$

Finally, if  $\mathcal{M}$  is Polish, then every  $\mu \in \mathfrak{P}(\mathcal{M})$  is also *tight* [52, Theorem 3.4.20], i.e.

$$\forall \mathcal{X} \in \mathcal{B}(\mathcal{M}) \quad \mu(\mathcal{X}) = \sup \{ \mu(\mathcal{K}) \mid \mathcal{K} \subseteq \mathcal{X}, \mathcal{K} \text{ compact} \}. \quad (4)$$

Borel probability measures with properties (3) and (4) are called *Radon probability measures*. Since we will be working with spacetimes (which are Polish spaces), all elements of  $\mathfrak{P}(\mathcal{M})$  will be Radon. For simplicity, from now on the term ‘measure’ will always stand for the ‘Borel probability measure’.

For any  $\mathcal{X} \subseteq \mathcal{M}$  its *indicator function*<sup>3</sup>  $\mathbf{1}_{\mathcal{X}} : \mathcal{M} \rightarrow \mathbb{R}$  is defined by  $\mathbf{1}_{\mathcal{X}}(p) = 1$  for  $p \in \mathcal{X}$  and  $\mathbf{1}_{\mathcal{X}}(p) = 0$  otherwise.  $\mathbf{1}_{\mathcal{X}}$  is a Borel function iff  $\mathcal{X} \in \mathcal{B}(\mathcal{M})$ .

A *simple function* on  $\mathcal{M}$  is any function  $s : \mathcal{M} \rightarrow \mathbb{R}$ , the range  $s(\mathcal{M})$  of which is finite. Such a function can be written in the form  $s = \sum_{i=1}^n \alpha_i \mathbf{1}_{\mathcal{X}_i}$  where  $s(\mathcal{M}) = \{\alpha_1, \dots, \alpha_n\}$  and  $\mathcal{X}_i = s^{-1}(\alpha_i)$  ( $i = 1, \dots, n$ ). Notice that  $s$  is Borel iff all  $\mathcal{X}_i$ s are Borel sets.

<sup>3</sup> The indicator function  $\mathbf{1}_{\mathcal{X}}$  is sometimes called the *characteristic function* of  $\mathcal{X}$ , but the latter term has another unrelated meaning in probability theory, which might cause confusion.

For any  $\mu \in \mathfrak{P}(\mathcal{M})$  the (Lebesgue) *integral* of a Borel nonnegative function  $f$  is defined as

$$\int_{\mathcal{M}} f d\mu := \sup \left\{ \sum_{i=1}^n \alpha_i \mu(\mathcal{X}_i) \mid \sum_{i=1}^n \alpha_i \mathbf{1}_{\mathcal{X}_i} \leq f \right\}.$$

It is well-defined by [51, Theorem 1.17], albeit it might be infinite. Now, for any Borel function  $f$  one introduces two nonnegative Borel functions  $f^{\pm} := \max\{\pm f, 0\}$  and defines  $\int_{\mathcal{M}} f d\mu := \int_{\mathcal{M}} f^+ d\mu - \int_{\mathcal{M}} f^- d\mu$  if at least one of the integrals is finite. For any  $\mathcal{X} \in \mathcal{B}(\mathcal{M})$  one additionally defines  $\int_{\mathcal{X}} f d\mu := \int_{\mathcal{M}} \mathbf{1}_{\mathcal{X}} f d\mu$ . A function  $f$  is  $\mu$ -integrable iff it is Borel and  $\int_{\mathcal{M}} |f| d\mu < +\infty$ . The space of  $\mu$ -integrable functions is denoted by  $\mathcal{L}^1(\mathcal{M}, \mu)$ . Observe that Borel bounded functions are  $\mu$ -integrable for any  $\mu \in \mathfrak{P}(\mathcal{M})$ .

We will often use the following classical theorem [51, Theorem 1.34].

**Theorem 2** (Lebesgue's dominated convergence theorem). *Let  $(f_n)_{n \in \mathbb{N}}$  be a sequence of Borel functions on  $\mathcal{M}$  such that  $f_n \rightarrow f$  pointwise. For any  $\mu \in \mathfrak{P}(\mathcal{M})$ , if there exists  $g \in \mathcal{L}^1(\mathcal{M}, \mu)$  such that  $|f_n| \leq g$  for all  $n \in \mathbb{N}$ , then also  $f \in \mathcal{L}^1(\mathcal{M}, \mu)$  and*

$$\lim_{n \rightarrow +\infty} \int_{\mathcal{M}} f_n d\mu = \int_{\mathcal{M}} f d\mu.$$

We will also make use of the following result [51, Theorem 1.39 (a)].

**Theorem 3.** *Let  $\mathcal{X} \in \mathcal{B}(\mathcal{M})$  and let  $f : \mathcal{M} \rightarrow \mathbb{R}$  be a Borel function such that  $f|_{\mathcal{X}} > 0$ . Then, for any  $\mu \in \mathfrak{P}(\mathcal{M})$ ,*

$$\int_{\mathcal{X}} f d\mu = 0 \quad \Leftrightarrow \quad \mu(\mathcal{X}) = 0.$$

Any Borel function  $f : \mathcal{M}_1 \rightarrow \mathcal{M}_2$  between topological spaces induces the *pushforward map*  $f_* : \mathfrak{P}(\mathcal{M}_1) \rightarrow \mathfrak{P}(\mathcal{M}_2)$ ,  $\mu \mapsto f_*\mu$ . The latter is called a *pushforward measure* and is defined through

$$\forall V \in \mathcal{B}(\mathcal{M}_2) \quad f_*\mu(V) := \mu(f^{-1}(V)).$$

As for the integrability, one has that  $g \in \mathcal{L}(\mathcal{M}_2, f_*\mu)$  iff  $g \circ f \in \mathcal{L}(\mathcal{M}_1, \mu)$ , in which case  $\int_{\mathcal{M}_2} g d(f_*\mu) = \int_{\mathcal{M}_1} g \circ f d\mu$ .

Given two probability spaces  $(\mathcal{M}_1, \mu_1), (\mathcal{M}_2, \mu_2)$ , there exists a unique measure  $\mu_1 \times \mu_2 \in \mathfrak{P}(\mathcal{M}_1 \times \mathcal{M}_2)$ , called the *product measure*, such that  $(\mu_1 \times \mu_2)(U_1 \times U_2) = \mu_1(U_1)\mu_2(U_2)$  for any  $U_i \in \mathcal{B}(\mathcal{M}_i)$ ,  $i = 1, 2$  (cf. [51] for details).

On the other hand, given  $\omega \in \mathfrak{P}(\mathcal{M}_1 \times \mathcal{M}_2)$ , its *marginals* are defined as  $(\text{pr}_i)_*\omega \in \mathfrak{P}(\mathcal{M}_i)$ , where  $\text{pr}_i : \mathcal{M}_1 \times \mathcal{M}_2 \rightarrow \mathcal{M}_i$  ( $i = 1, 2$ ) are the canonical projection maps. Obviously, the marginals of the product measure  $\mu_1 \times \mu_2$  are  $\mu_1$  and  $\mu_2$ ; however, usually there are many measures on  $\mathcal{M}_1 \times \mathcal{M}_2$  sharing the same pair of marginals.

Given a measure  $\mu \in \mathfrak{P}(\mathcal{M})$ , its *support* can be defined as the smallest closed set with full measure. Symbolically,

$$\text{supp } \mu := \bigcap \{F \subseteq \mathcal{M} \mid F \text{ closed, } \mu(F) = 1\}.$$

### 2.3. Causality Theory

For a detailed exposition of causality theory the reader is referred to [7, 43, 46, 48].

Recall that a *spacetime* is a connected time-oriented Lorentzian manifold. Causality theory introduces and studies certain binary relations between points (i.e. *events*) of a given spacetime  $\mathcal{M}$ . Namely, for any  $p, q \in \mathcal{M}$ , we say that  $p$  *causally* (*chronologically*) *precedes*  $q$ , what is denoted by  $p \preceq q$  (resp.  $p \ll q$ ), iff there exists a piecewise smooth future-directed causal (resp. timelike) curve  $\gamma : [0, 1] \rightarrow \mathcal{M}$  from  $p$  to  $q$ , i.e.  $\gamma(0) = p$  and  $\gamma(1) = q$ . Additionally, we say that  $p$  *horismotically precedes*  $q$ , what is denoted by  $p \rightarrow q$ , iff  $p \preceq q$ , but  $p \not\ll q$ .

Clearly, the relations  $\preceq$  and  $\ll$  are transitive and  $\preceq$  is also reflexive. Moreover ([46, Chapter 14, Corollary 1]),

$$\forall p, q, r \in \mathcal{M} \quad p \ll r \preceq q \vee p \preceq r \ll q \Rightarrow p \ll q. \quad (5)$$

To denote  $\preceq$  ( $\ll$ ,  $\rightarrow$ ) understood as a subset of  $\mathcal{M}^2$  it is customary to use the symbol  $J^+$  (resp.  $I^+$ ,  $E^+$ ).  $I^+$  is open and equal to  $\text{int } J^+$ , and so the *causal structure* of  $\mathcal{M}$  is completely determined by the relation  $\preceq$  and the topology of  $\mathcal{M}$ . Moreover,  $\bar{I}^+ = \bar{J}^+$ ,  $\partial I^+ = \partial J^+$  and  $E^+ = J^+ \setminus I^+$ .

For any  $\mathcal{X} \subseteq \mathcal{M}$  one defines

$$J^+(\mathcal{X}) := \text{pr}_2((\mathcal{X} \times \mathcal{M}) \cap J^+) \quad \text{and} \quad J^-(\mathcal{X}) := \text{pr}_1((\mathcal{M} \times \mathcal{X}) \cap J^+). \quad (6)$$

If  $\mathcal{X}$  is a singleton, one simply writes  $J^\pm(p)$  instead of  $J^\pm(\{p\})$ . Notice that  $J^\pm(\mathcal{X}) = \bigcup_{p \in \mathcal{X}} J^\pm(p)$ .

Let now  $U \subseteq \mathcal{M}$  be an open subset of  $\mathcal{M}$ . One defines  $\preceq_U$  to be the causal precedence relation on  $U$  treated as a spacetime on its own right. By analogy with  $J^+$ , we denote  $J_U^+ := \{(p, q) \in U^2 \mid p \preceq_U q\}$ . Notice that  $J_U^+ \subseteq J^+ \cap U^2$ , but not necessarily *vice versa*, because  $p \preceq_U q$  requires a piecewise smooth future-directed causal curve from  $p$  to  $q$  not only to exist, but also to be contained in  $U$ .

Analogously to (6), one defines  $J_U^\pm(\mathcal{X})$  for any subset  $\mathcal{X} \subseteq \mathcal{M}$ .

One similarly introduces  $I^\pm(\mathcal{X})$ ,  $I_U^+$ ,  $I_U^\pm(\mathcal{X})$  and  $E^\pm(\mathcal{X})$ ,  $E_U^+$ ,  $E_U^\pm(\mathcal{X})$ . Observe that, by (5),  $J^+(\mathcal{X}) = I^+(\mathcal{X})$  for any open  $\mathcal{X} \subseteq \mathcal{M}$ .

A subset  $\mathcal{F} \subseteq \mathcal{M}$  is called a *future set* iff<sup>4</sup>  $J^+(\mathcal{F}) = \mathcal{F}$ . Similarly, subset  $\mathcal{P} \subseteq \mathcal{M}$  is called a *past set* iff  $J^-(\mathcal{P}) = \mathcal{P}$ . Usually, one requires the future and past sets to be open by definition. However, if we drop this assumption future and past sets behave more naturally under set-theoretical operations.

**Proposition 1.**  $\mathcal{F} \subseteq \mathcal{M}$  is a future set iff  $\mathcal{F}^c$  is a past set.

<sup>4</sup> Notice that only the inclusion ' $\subseteq$ ' is nontrivial in the definition of a future (past) set.



*Proof.* The statement is proven by the following chain of equivalences:

$$\begin{aligned}
 J^-(\mathcal{F}^c) \subseteq \mathcal{F}^c &\Leftrightarrow \forall s \in \mathcal{M} \quad [\exists r \in \mathcal{F}^c \quad s \preceq r] \Rightarrow s \in \mathcal{F}^c \\
 &\Leftrightarrow \forall s \in \mathcal{M} \quad [\exists r \in J^+(s) \setminus \mathcal{F}] \Rightarrow s \notin \mathcal{F} \\
 &\Leftrightarrow \forall s \in \mathcal{M} \quad J^+(s) \setminus \mathcal{F} = \emptyset \Leftarrow s \in \mathcal{F} \\
 &\Leftrightarrow \forall s \in \mathcal{F} \quad J^+(s) \subseteq \mathcal{F} \\
 &\Leftrightarrow \bigcup_{s \in \mathcal{F}} J^+(s) \subseteq \mathcal{F} \Leftrightarrow J^+(\mathcal{F}) \subseteq \mathcal{F}. \quad \square
 \end{aligned}$$

**Proposition 2.** Let  $\{\mathcal{X}_\alpha\}_{\alpha \in A}$  be a family of future (past) subsets of  $\mathcal{M}$ . Then, also  $\bigcup_{\alpha \in A} \mathcal{X}_\alpha$  and  $\bigcap_{\alpha \in A} \mathcal{X}_\alpha$  are future (past) subsets of  $\mathcal{M}$ .

*Proof.* Assuming that all  $\mathcal{X}_\alpha$ s are future sets, notice that  $J^+(\bigcup_{\alpha \in A} \mathcal{X}_\alpha) = \bigcup_{\alpha \in A} J^+(\mathcal{X}_\alpha) = \bigcup_{\alpha \in A} \mathcal{X}_\alpha$ . If  $\mathcal{X}_\alpha$ s are past sets, simply replace  $J^+$  with  $J^-$  in the previous sentence.

We have thus shown that a union of future (past) sets is a future (past) set. To obtain an analogous result for the intersection, one simply uses Proposition 1 and de Morgan's laws.  $\square$

A function  $f : \mathcal{M} \rightarrow \mathbb{R}$  is called

- a *causal function* iff it is nondecreasing along every future-directed causal curve;
- a *generalised time function* iff it is increasing along every future-directed causal curve;
- a *time function* iff it is a continuous generalised time function;
- a *temporal function* iff it is a smooth function with past-directed timelike gradient.

Each of the above properties is stronger than the preceding one.

Causal functions can be characterised by means of future sets.

**Proposition 3.** Let  $\mathcal{M}$  be a spacetime. For any function  $f : \mathcal{M} \rightarrow \mathbb{R}$  the following conditions are equivalent

- (i)  $f$  is causal,
- (ii)  $f^{-1}((a, +\infty))$  is a future set for any  $a \in \mathbb{R}$ ,
- (iii)  $f^{-1}([a, +\infty))$  is a future set for any  $a \in \mathbb{R}$ .

*Proof.* ‘(i)  $\Rightarrow$  (ii)’ Assume that  $f$  is causal and  $a \in \mathbb{R}$ . If  $f^{-1}((a, +\infty)) = \emptyset$ , then it is trivially a future set. Suppose then that  $f^{-1}((a, +\infty)) \neq \emptyset$  and take any  $q \in J^+(f^{-1}((a, +\infty)))$ , which means that there exists  $p \preceq q$  such that  $f(p) > a$ . By the causality of  $f$  we have that  $f(q) \geq f(p) > a$  and so  $q \in f^{-1}((a, +\infty))$ . We thus obtain the inclusion  $J^+(f^{-1}((a, +\infty))) \subseteq f^{-1}((a, +\infty))$ . The other inclusion is obvious.

‘(ii)  $\Rightarrow$  (iii)’ Observe that  $f^{-1}([a, +\infty)) = \bigcap_{n=1}^{\infty} f^{-1}((a - \frac{1}{n}, +\infty))$ . By (ii) and Proposition 2, we obtain (iii).

‘(iii)  $\Rightarrow$  (i)’ Assume  $f$  is not causal, i.e. there exist  $p, q \in \mathcal{M}$  such that  $p \preceq q$  but  $f(p) > f(q)$ . We claim that  $f^{-1}([f(p), +\infty))$  is not a future set. Indeed, were it a future set, then, since it clearly contains  $p$ , it would contain

$q$  as well. But this would mean that  $f(q) \geq f(p)$ , in contradiction with the assumption.  $\square$

On the other hand, future sets can be characterised by means of their indicator functions.

**Corollary 1.** *Let  $\mathcal{M}$  be a spacetime.  $\mathcal{F} \subseteq \mathcal{M}$  is a future set iff the function  $\mathbf{1}_{\mathcal{F}}$  is causal.*

*Proof.* Observe that

$$\mathbf{1}_{\mathcal{F}}^{-1}([a, +\infty)) = \begin{cases} \mathcal{M} & \text{for } a \leq 0 \\ \mathcal{F} & \text{for } 0 < a \leq 1 \\ \emptyset & \text{for } a > 1 \end{cases}.$$

By the equivalence ‘(i)  $\Leftrightarrow$  (iii)’ from Proposition 3, we immediately obtain the desired result.  $\square$

An *admissible measure* on  $\mathcal{M}$  is any  $\eta \in \mathfrak{P}(\mathcal{M})$  such that ([7, Definition 3.19])

- for any nonempty open subset  $U \subseteq \mathcal{M}$   $\eta(U) > 0$ ,
- for any  $p \in \mathcal{M}$  the boundaries  $\partial I^{\pm}(p)$  are  $\eta$ -null.

To such an  $\eta$  one associates the functions  $t^{-}, t^{+} : \mathcal{M} \rightarrow \mathbb{R}$ , called *past* and *future* volume functions, respectively, defined via

$$\forall p \in \mathcal{M} \quad t^{\pm}(p) := \mp \eta(I^{\pm}(p)).$$

Volume functions are causal and semi-continuous, and hence Borel.

For any  $p, q \in \mathcal{M}$  let  $\hat{C}(p, q)$  denote the set of piecewise smooth future-directed causal curves from  $p$  to  $q$ . The *Lorentzian distance* (or *time separation*) is the map  $d : \mathcal{M}^2 \rightarrow [0, +\infty]$  defined by

$$d(p, q) := \begin{cases} \sup_{\gamma \in \hat{C}(p, q)} \int_0^1 \sqrt{-g_{\alpha\beta} \dot{\gamma}^{\alpha} \dot{\gamma}^{\beta}} dt & \text{if } \hat{C}(p, q) \neq \emptyset \\ 0 & \text{if } \hat{C}(p, q) = \emptyset \end{cases}.$$

Its basic properties include:

- For any  $p, q \in \mathcal{M}$   $d(p, q) > 0 \Leftrightarrow p \ll q$ .
- The *reverse triangle inequality* holds. Namely, for any  $p, q, r \in \mathcal{M}$

$$p \preceq r \preceq q \Rightarrow d(p, r) + d(r, q) \leq d(p, q). \quad (7)$$

- If there exists a timelike loop through  $p \in \mathcal{M}$  (i.e. a piecewise smooth curve from  $p$  to  $p$ ), then  $d(p, p) = +\infty$ . Otherwise  $d(p, p) = 0$ .
- For any  $p, q \in \mathcal{M}$ , if  $d(p, q) \in (0, +\infty)$  then  $d(q, p) = 0$ .
- The map  $d$  is lower semi-continuous [46, Chapter 14, Lemma 17] and hence Borel.

The *causal ladder* is a hierarchy of spacetimes according to strictly increasing requirements on their causal properties [7]. The rungs of this ladder, from the top to the bottom, read:

$$\text{Globally hyperbolic} \Rightarrow \text{Causally simple} \Rightarrow \text{Causally continuous}$$

$\Rightarrow$  Stably causal  $\Rightarrow$  Strongly causal  $\Rightarrow$  Distinguishing  
 $\Rightarrow$  Causal  $\Rightarrow$  Chronological.

Each level of the hierarchy can be defined in many equivalent ways. Below we present only these definitions, characterisations and properties, of which we make use in the paper. For a complete review of the causal hierarchy, consult [43, Section 3].

$\mathcal{M}$  is *chronological* iff it satisfies one of the following equivalent conditions:

- (i)  $p \not\ll p$  for all  $p \in \mathcal{M}$ .
- (ii) No timelike loop exists.
- (iii) Any volume function is increasing along every future-directed timelike curve.
- (iv)  $d(p, p) = 0$  for all  $p \in \mathcal{M}$ .

$\mathcal{M}$  is *causal* iff it satisfies one of the following equivalent conditions:

- (i) The relation  $\preceq$  is a *partial order*; hence, in addition to being reflexive and transitive, it is also antisymmetric.
- (ii) No causal loop exists.

$\mathcal{M}$  is *future (past) distinguishing* iff it satisfies one of the following equivalent conditions:

- (i) For any  $p, q \in \mathcal{M}$ , the equality  $I^+(p) = I^+(q)$  (resp.  $I^-(p) = I^-(q)$ ) implies that  $p = q$ .
- (ii) Any future (past) volume function is a generalised time function [7, Proposition 3.24].

$\mathcal{M}$  is *distinguishing* iff it is both future and past distinguishing.

$\mathcal{M}$  is *strongly causal* iff the family  $\{I^+(p) \cap I^-(q) \mid p, q \in \mathcal{M}\}$  is a base of the standard manifold topology of  $\mathcal{M}$ . It is *stably causal* iff it admits a time function or, equivalently, iff it admits a temporal function [8]. It is *causally continuous* iff any volume function is a time function.

$\mathcal{M}$  is *causally simple* iff it is causal and satisfies one of the following equivalent conditions [43, Proposition 3.68]:

- (i)  $J^+(p)$  and  $J^-(p)$  are closed for every  $p \in \mathcal{M}$ ;
- (ii)  $J^+(\mathcal{K})$  and  $J^-(\mathcal{K})$  are closed for every compact  $\mathcal{K} \subseteq \mathcal{M}$ ;
- (iii)  $J^+$  is a closed subset of  $\mathcal{M}^2$ .

Before providing a definition of the top level of the causal hierarchy, recall that a curve  $\gamma : (a, b) \rightarrow \mathcal{M}$  with  $-\infty \leq a < b \leq +\infty$  is called *extendible* iff it has a continuous extension onto  $[a, b)$  or onto  $(a, b]$ . Otherwise such a curve is called *inextendible*. Recall also that a *Cauchy hypersurface* is a subset  $\mathcal{S} \subseteq \mathcal{M}$  which is met exactly once by any inextendible timelike curve. Any such  $\mathcal{S}$  is a connected, closed, *achronal* (i.e.  $\mathcal{S}^2 \cap I^+ = \emptyset$ ) topological hypersurface, met by every inextendible causal curve [46, Chapter 14, Lemma 29 & Proposition 31]. However, such an  $\mathcal{S}$  need not be *acausal* (i.e.  $\mathcal{S}^2 \cap J^+$  might be nonempty).

$\mathcal{M}$  is *globally hyperbolic* iff it satisfies one of the following equivalent conditions:

- (i)  $\mathcal{M}$  is causal and the sets  $J^+(p) \cap J^-(q)$  are compact for all  $p, q \in \mathcal{M}$ ;
- (ii)  $\mathcal{M}$  admits a smooth temporal function  $\mathcal{T}$ , the level sets of which are (smooth spacelike) Cauchy hypersurfaces [8].

In a globally hyperbolic spacetime the Lorentzian distance  $d$  is finite-valued and continuous. Moreover, for every  $(p, q) \in J^+$  there exists a causal geodesic  $\gamma$  from  $p$  to  $q$  of length  $d(p, q)$  [46, Chapter 14].

### 3. On the $\sigma$ -Compactness of $J^+$

The purpose of this section is to prove the following theorem.

**Theorem 4.** *Let  $\mathcal{M}$  be a spacetime. Then,  $J^+ \subseteq \mathcal{M}^2$  is a  $\sigma$ -compact set.*

Let us note here that this property is automatic in causally simple spacetimes. Indeed, let  $(K_n)_{n \in \mathbb{N}}$  be an exhaustion of  $\mathcal{M}$  with compact sets and notice that  $J^+ = \bigcup_{m, n \in \mathbb{N}} J^+ \cap (K_n \times K_m)$ . But  $J^+ \subseteq \mathcal{M}^2$  is a *closed* subset for  $\mathcal{M}$  causally simple, therefore  $J^+ \cap (K_n \times K_m)$  is compact for any  $m, n \in \mathbb{N}$ .

In the proof of Theorem 4, however, we shall make no assumptions on the causal properties of  $\mathcal{M}$ .

Theorem 4 implies that  $J^+$  is Borel for any spacetime. As we shall see, it also implies that  $J^\pm(\mathcal{X})$  is Borel for any closed  $\mathcal{X} \subseteq \mathcal{M}$ . Moreover, previous statements are still true if we replace  $J^\pm$  with  $E^\pm$ .

Theorem 4 is thus settled in the overlap of causality theory, topology and measure theory. Whereas the interplay between the causal and topological properties of spacetimes is relatively well understood, the questions concerning Borelness have never been, to authors' best knowledge, addressed in the relativistic context. The study of the interaction between causality and measure theory is, however, essential from the viewpoint of the theory developed in Sect. 4.

We recall the notion of *simple convex sets* (called also *simple regions*) [48, Section 1]. Loosely speaking, they are small patches of the spacetime  $\mathcal{M}$  with 'nice' topological, differential and causal properties, which constitute a countable cover of the entire spacetime.

Concretely, let  $\mathcal{M}$  be a spacetime. Then, for any  $p \in \mathcal{M}$  there exists a star-shaped neighbourhood  $Q \subseteq T_p\mathcal{M}$  containing the zero vector and such that the exponential map  $\exp_p$  restricted to  $Q$  is a diffeomorphism. The image of this diffeomorphism  $\exp_p(Q)$  is called a *normal neighbourhood* of  $p$ . Every event has a neighbourhood  $U$  which is a normal neighbourhood of any  $p \in U$ . Such a  $U$  is called *convex*. If  $U \subseteq \mathcal{M}$  is convex, then it is open and for any  $p, q \in U$  there exists precisely one geodesic from  $p$  to  $q$  which is contained in  $U$  [46, p. 129].

From the point of view of causality theory, the following property of convex sets will be crucial: if  $U \subseteq \mathcal{M}$  is convex, then  $J_U^+$  is a closed subset of  $U^2$  [46, Lemma 14.2].

Finally, a convex set  $N$  is called *simple* iff  $\overline{N}$  is compact and contained in another convex set  $U$ .

Any spacetime  $\mathcal{M}$  can be covered with a family of simple convex sets [48, Proposition 1.13]. This cover can be chosen countable, because every spacetime is a Lindelöf space.

*Proof of Theorem 4.* Fix a countable, locally finite family of simple convex sets  $\{N_i\}_{i \in \mathbb{N}}$  covering  $\mathcal{M}$ . Let also  $\{U_i\}_{i \in \mathbb{N}}$  be a family of convex sets such that  $\forall i \in \mathbb{N} \ \overline{N}_i \subseteq U_i$ , which exists by the very definition of a simple convex set.

We introduce a couple more definitions.

Take any  $i \in \mathbb{N}$ . Recall that  $J_{U_i}^+$  is a closed subset of  $U_i^2$ , whereas  $\overline{N}_i^2$  is a compact subset of  $U_i^2$ . Let us first define the following compact subset of  $U_i^2$

$$J_{(\overline{N}_i)}^+ := J_{U_i}^+ \cap \overline{N}_i^2 = \left\{ (p, q) \in \overline{N}_i^2 \mid p \preceq_{U_i} q \right\},$$

that is the set containing all these pairs of points from  $\overline{N}_i$  which can be connected by a piecewise smooth future-directed causal curve contained in  $U_i$ . For any  $\mathcal{X} \subseteq \mathcal{M}$  define, by analogy with (6),

$$\begin{aligned} J_{(\overline{N}_i)}^+(\mathcal{X}) &:= \text{pr}_2 \left( (\mathcal{X} \times \mathcal{M}) \cap J_{(\overline{N}_i)}^+ \right) \quad \text{and} \\ J_{(\overline{N}_i)}^-(\mathcal{X}) &:= \text{pr}_1 \left( (\mathcal{M} \times \mathcal{X}) \cap J_{(\overline{N}_i)}^+ \right). \end{aligned} \quad (8)$$

If  $\mathcal{X}$  is a singleton, one simply writes  $J_{(\overline{N}_i)}^\pm(p)$  instead of  $J_{(\overline{N}_i)}^\pm(\{p\})$ . Notice that if  $\mathcal{X}$  is closed, then  $J_{(\overline{N}_i)}^\pm(\mathcal{X})$  is a compact subset of  $U_i$ .

For the next definition, fix  $i_1, i_2 \in \mathbb{N}$  and introduce

$$\begin{aligned} J_{(\overline{N}_{i_1}, \overline{N}_{i_2})}^+ &:= \left\{ (p, q) \in \overline{N}_{i_1} \times \overline{N}_{i_2} \mid \exists r \in \overline{N}_{i_1} \cap \overline{N}_{i_2} \ p \preceq_{U_{i_1}} r \preceq_{U_{i_2}} q \right\} \\ &= \bigcup_{r \in \overline{N}_{i_1} \cap \overline{N}_{i_2}} J_{\overline{N}_{i_1}}^-(r) \times J_{\overline{N}_{i_2}}^+(r). \end{aligned}$$

This is the set of all those pairs of points  $(p, q) \in \overline{N}_{i_1} \times \overline{N}_{i_2}$ , which can be connected by a concatenation of two piecewise smooth future-directed causal curves, first of which is contained in  $U_{i_1}$ , while the other in  $U_{i_2}$ , and the concatenation point  $r$  must lie in the compact set  $\overline{N}_{i_1} \cap \overline{N}_{i_2}$ . As above, we additionally define, for any  $\mathcal{X} \subseteq \mathcal{M}$ ,

$$\begin{aligned} J_{(\overline{N}_{i_1}, \overline{N}_{i_2})}^+(\mathcal{X}) &:= \text{pr}_2 \left( (\mathcal{X} \times \mathcal{M}) \cap J_{(\overline{N}_{i_1}, \overline{N}_{i_2})}^+ \right) \quad \text{and} \\ J_{(\overline{N}_{i_1}, \overline{N}_{i_2})}^-(\mathcal{X}) &:= \text{pr}_1 \left( (\mathcal{M} \times \mathcal{X}) \cap J_{(\overline{N}_{i_1}, \overline{N}_{i_2})}^+ \right). \end{aligned} \quad (9)$$

Finally, fix  $n \geq 3$  together with  $i_1, i_2, \dots, i_n \in \mathbb{N}$  and define, recursively,

$$\begin{aligned} J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^+ &:= \left\{ (p, q) \in \overline{N}_{i_1} \times \overline{N}_{i_n} \mid \exists r \in \overline{N}_{i_{n-1}} \ (p, r) \in J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_{n-1}})}^+ \right. \\ &\quad \left. \wedge \ (r, q) \in J_{(\overline{N}_{i_{n-1}}, \overline{N}_{i_n})}^+ \right\} \end{aligned}$$



$$\begin{aligned}
J_{(\overline{N}_{i_1}, \overline{N}_{i_2})}^+ &= \bigcap_{k \in \mathbb{N}} \bigcup_{m \in F_k} J_{(\overline{N}_{i_1})}^- (\overline{B}(a_m, \frac{1}{k})) \times J_{(\overline{N}_{i_2})}^+ (\overline{B}(a_m, \frac{1}{k})) \\
&= \left\{ (p, q) \in \overline{N}_{i_1} \times \overline{N}_{i_2} \mid \forall k \in \mathbb{N} \exists m \in F_k \exists p_m, q_m \in \overline{B}(a_m, \frac{1}{k}) \right. \\
&\quad \left. p \preceq_{U_{i_1}} p_m \wedge q_m \preceq_{U_{i_2}} q \right\}, \tag{11}
\end{aligned}$$

which would mean that  $J_{(\overline{N}_{i_1}, \overline{N}_{i_2})}^+$  is a closed subset of  $\overline{N}_{i_1} \times \overline{N}_{i_2}$  (hence also a compact subset of  $U_{i_1} \times U_{i_2}$ ), because finite unions of closed sets are closed and so are any intersections of closed sets.

Indeed, to prove the inclusion ' $\subseteq$ ', assume  $(p, q) \in \overline{N}_{i_1} \times \overline{N}_{i_2}$  is such that there exists  $r \in \overline{N}_{i_1} \cap \overline{N}_{i_2}$  satisfying  $p \preceq_{U_{i_1}} r \preceq_{U_{i_2}} q$ . For any  $k \in \mathbb{N}$ , since  $\{B(a_m, \frac{1}{k})\}_{m \in F_k}$  covers  $\overline{N}_{i_1} \cap \overline{N}_{i_2}$ , it is possible to find  $m \in F_k$  such that  $r \in B(a_m, \frac{1}{k})$ . One can thus take  $p_m := r =: q_m$ .

On the other hand, to show the inclusion ' $\supseteq$ ', let us assume that  $(p, q) \in \overline{N}_{i_1} \times \overline{N}_{i_2}$  is such that

$$\forall k \in \mathbb{N} \exists m \in F_k \exists p_m, q_m \in \overline{B}(a_m, \frac{1}{k}) \quad p \preceq_{U_{i_1}} p_m \text{ and } q_m \preceq_{U_{i_2}} q.$$

We can thus construct the sequence  $(a_{m_k})_{k \in \mathbb{N}}$ , which, being contained in the compact set  $\overline{N}_{i_1} \cap \overline{N}_{i_2}$ , has a subsequence  $(a_{m_{k_l}})_{l \in \mathbb{N}}$  convergent to some  $a_\infty \in \overline{N}_{i_1} \cap \overline{N}_{i_2}$ . Notice now that because  $p_{m_k}, q_{m_k} \in \overline{B}(a_{m_k}, \frac{1}{k})$  for any  $k \in \mathbb{N}$ , therefore also

$$\lim_{l \rightarrow +\infty} p_{m_{k_l}} = \lim_{l \rightarrow +\infty} q_{m_{k_l}} = a_\infty.$$

We now invoke the fact that  $J_{U_{i_1}}^+$  and  $J_{U_{i_2}}^+$  are *closed* subsets of  $U_{i_1}^2$  and of  $U_{i_2}^2$ , respectively. It implies that

$$p \preceq_{U_{i_1}} a_\infty \preceq_{U_{i_2}} q,$$

which completes the proof of (11) and of the base case of the induction.

We now move to the proof of the inductive step, which essentially goes along the same lines as the proof of the base case.

The assumption says that for any  $i_1, \dots, i_n \in \mathbb{N}$ ,  $J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^+$  is a compact subset of  $U_{i_1} \times U_{i_n}$ . The induction hypothesis then reads: for any  $i_1, \dots, i_{n+1} \in \mathbb{N}$ ,  $J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_{n+1}})}^+$  is a compact subset of  $U_{i_1} \times U_{i_{n+1}}$ .

Let now  $\{a_m\}_{m \in \mathbb{N}}$  be a dense subset of  $N_{i_n}$ , hence also a dense subset of  $\overline{N}_{i_n}$ . Similarly as before, for each  $k \in \mathbb{N}$  consider the family  $\{B(a_m, \frac{1}{k})\}_{m \in \mathbb{N}}$  covering  $\overline{N}_{i_n}$ , and take its finite subcover  $\{B(a_m, \frac{1}{k})\}_{m \in F_k}$ .

We now claim that

$$\begin{aligned}
&J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_{n+1}})}^+ \\
&= \bigcap_{k \in \mathbb{N}} \bigcup_{m \in F_k} J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^- (\overline{B}(a_m, \frac{1}{k})) \times J_{(\overline{N}_{i_n}, \overline{N}_{i_{n+1}})}^+ (\overline{B}(a_m, \frac{1}{k})) \\
&= \left\{ (p, q) \in \overline{N}_{i_1} \times \overline{N}_{i_{n+1}} \mid \forall k \in \mathbb{N} \exists m \in F_k \exists p_m, q_m \in \overline{B}(a_m, \frac{1}{k}) \right. \\
&\quad \left. (p, p_m) \in J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^+ \wedge (q_m, q) \in J_{(\overline{N}_{i_n}, \overline{N}_{i_{n+1}})}^+ \right\}, \tag{12}
\end{aligned}$$

which would mean that  $J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_{n+1}})}^+$  is a closed subset of  $\overline{N}_{i_1} \times \overline{N}_{i_{n+1}}$  (hence also a compact subset of  $U_{i_1} \times U_{i_{n+1}}$ ), because we already know that  $J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^-(\overline{B}(a_m, \frac{1}{k}))$  is closed in  $\overline{N}_{i_1}$  (by the induction assumption and definitions (10)) and that  $J_{(\overline{N}_{i_n}, \overline{N}_{i_{n+1}})}^+(\overline{B}(a_m, \frac{1}{k}))$  is closed in  $\overline{N}_{i_{n+1}}$  (by the base case and definitions (9)).

To show the inclusion ‘ $\subseteq$ ’ in (12), assume  $(p, q) \in \overline{N}_{i_1} \times \overline{N}_{i_{n+1}}$  is such that there exists  $r \in \overline{N}_{i_n}$  satisfying  $(p, r) \in J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^+$  and  $(r, q) \in J_{(\overline{N}_{i_n}, \overline{N}_{i_{n+1}})}^+$ . For any  $k \in \mathbb{N}$ , since  $\{B(a_m, \frac{1}{k})\}_{m \in F_k}$  covers  $\overline{N}_{i_n}$ , it is possible to find  $m \in F_k$  such that  $r \in B(a_m, \frac{1}{k})$ . One can thus take  $p_m := r =: q_m$ .

On the other hand, to show the inclusion ‘ $\supseteq$ ’, let us assume that  $(p, q) \in \overline{N}_{i_1} \times \overline{N}_{i_{n+1}}$  are such that

$$\begin{aligned} \forall k \in \mathbb{N} \quad \exists m \in F_k \quad \exists p_m, q_m \in \overline{B}(a_m, \frac{1}{k}) \\ (p, p_m) \in J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^+, \quad (q_m, q) \in J_{(\overline{N}_{i_n}, \overline{N}_{i_{n+1}})}^+. \end{aligned}$$

We can thus construct a sequence  $(a_{m_k})_{k \in \mathbb{N}}$ , which, being contained in the compact set  $\overline{N}_{i_n}$ , has a subsequence  $(a_{m_{k_l}})_{l \in \mathbb{N}}$  convergent to some  $a_\infty \in \overline{N}_{i_n}$ . Analogously as before, we argue that also the sequences  $(p_{m_k}), (q_{m_k})$  have subsequences converging to  $a_\infty$ .

By the induction assumption, we obtain that  $(p, a_\infty) \in J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^+$ . On the other hand, invoking the base case we similarly obtain that  $(a_\infty, q) \in J_{(\overline{N}_{i_n}, \overline{N}_{i_{n+1}})}^+$ . This completes the proof of (12) and of the entire induction.

Altogether, we can thus write that

$$\begin{aligned} \forall n \in \mathbb{N} \quad \forall i_1, i_2, \dots, i_n \in \mathbb{N} \\ J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^+ \text{ is a compact subset of } U_{i_1} \times U_{i_n}, \text{ and hence of } \mathcal{M}^2. \end{aligned} \quad (13)$$

Bearing the above in mind, the  $\sigma$ -compactness of  $J^+$  will be proven if we show that

$$J^+ = \bigcup_{n=1}^{\infty} \bigcup_{i_1, i_2, \dots, i_n \in \mathbb{N}} J_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}^+. \quad (14)$$

In order to show the inclusion ‘ $\subseteq$ ’, take any  $(p, q) \in J^+$  and let  $\gamma : [0, 1] \rightarrow \mathcal{M}$  be a piecewise smooth future-directed causal curve from  $p$  to  $q$ .

Consider the inverse images  $\gamma^{-1}(N_i)$ ,  $i \in \mathbb{N}$ . By the continuity of  $\gamma$ , they are all open subsets of  $[0, 1]$ ; however, they might be *disconnected* (i.e. they need not be intervals). Nevertheless, every  $\gamma^{-1}(N_i)$  is a union of its connected components, which are all open<sup>5</sup> subintervals of  $[0, 1]$ .

Let us thus consider the family of all connected components of all  $\gamma^{-1}(N_i)$ ’s,  $i \in \mathbb{N}$ . This family is a cover of  $[0, 1]$ , and since the latter is a compact space, we can take its finite subcover  $\mathcal{I} := \{I_1, I_2, \dots, I_n\}$ , where each of

<sup>5</sup> Locally compact spaces (and  $[0, 1]$  is such a space) can be characterised as the spaces in which every connected component of every open set is itself open.



the intervals  $I_j$  ( $j = 1, \dots, n$ ) is a connected component of some (possibly not unique)  $\gamma^{-1}(N_{i_j})$ . Therefore

$$\forall j = 1, \dots, n \quad \gamma(I_j) \subseteq N_{i_j}$$

and, by the continuity of  $\gamma$ ,

$$\forall j = 1, \dots, n \quad \gamma(\bar{I}_j) \subseteq \bar{N}_{i_j}.$$

Without loss of generality, we can assume that  $I_{j_1} \not\subseteq I_{j_2}$  for all  $j_1 \neq j_2$ . Bearing this in mind, we can rewrite  $\mathcal{I}$  either as  $\{[0, 1]\}$  (the trivial cover) or, if  $n > 1$ , as

$$\mathcal{I} = \{[0, b_1], (a_2, b_2), \dots, (a_{n-1}, b_{n-1}), (a_n, 1]\},$$

where  $0 < a_2 < a_3 < \dots < a_n < 1$ . Notice also that  $b_j > a_{j+1}$  for  $j = 1, \dots, n-1$ , because otherwise such an  $\mathcal{I}$  would not be a cover.

In the first (trivial) case,  $\gamma([0, 1]) \subseteq N_{i_1} \subseteq \bar{N}_{i_1}$  for some  $i_1 \in \mathbb{N}$  and hence  $(p, q) \in J_{(\bar{N}_{i_1})}^+$ .

In the second case, observe that

$$\begin{aligned} \gamma([0, a_2]) &\subseteq \gamma([0, b_1]) \subseteq N_{i_1} \subseteq \bar{N}_{i_1}, \\ \gamma([a_2, a_3]) &\subseteq \gamma([a_2, b_2]) \subseteq \bar{N}_{i_2}, \\ &\dots \\ \gamma([a_j, a_{j+1}]) &\subseteq \gamma([a_j, b_j]) \subseteq \bar{N}_{i_j}, \\ &\dots \\ \gamma([a_{n-1}, a_n]) &\subseteq \gamma([a_{n-1}, b_{n-1}]) \subseteq \bar{N}_{i_{n-1}}, \\ \gamma([a_n, 1]) &\subseteq \bar{N}_{i_n}, \end{aligned}$$

for some  $i_1, \dots, i_n \in \mathbb{N}$  and hence  $(p, q) \in J_{(\bar{N}_{i_1}, \dots, \bar{N}_{i_n})}^+$ .

In either case, we obtain that  $(p, q) \in \bigcup_{n=1}^{\infty} \bigcup_{i_1, i_2, \dots, i_n \in \mathbb{N}} J_{(\bar{N}_{i_1}, \bar{N}_{i_2}, \dots, \bar{N}_{i_n})}^+$ .

In order to show the other inclusion ‘ $\supseteq$ ’ in (14), notice simply that a concatenation of finitely many piecewise smooth future-directed causal curves is itself a piecewise smooth future-directed causal curve. Therefore, if  $(p, q) \in J_{(\bar{N}_{i_1}, \bar{N}_{i_2}, \dots, \bar{N}_{i_n})}^+$ , then  $(p, q) \in J^+$ .  $\square$

**Corollary 2.** *Let  $\mathcal{M}$  be a spacetime. Then,  $E^+$  is a  $\sigma$ -compact subset of  $\mathcal{M}^2$ .*

*Proof.* On the strength of (14), we have that

$$E^+ := J^+ \setminus I^+ = \bigcup_{n=1}^{\infty} \bigcup_{i_1, i_2, \dots, i_n \in \mathbb{N}} J_{(\bar{N}_{i_1}, \bar{N}_{i_2}, \dots, \bar{N}_{i_n})}^+ \setminus I^+$$

and since  $I^+$  is an open subset of  $\mathcal{M}^2$ ,  $J_{(\bar{N}_{i_1}, \bar{N}_{i_2}, \dots, \bar{N}_{i_n})}^+ \setminus I^+$  is a closed subset of  $\bar{N}_{i_1} \times \bar{N}_{i_n}$  (for any  $i_1, i_2, \dots, i_n \in \mathbb{N}$ ), and hence a compact subset of  $\mathcal{M}^2$ .  $\square$

**Corollary 3.** *Let  $\mathcal{M}$  be a spacetime and let  $\mathcal{X} \subseteq \mathcal{M}$  be a countable union of closed sets. Then,  $J^\pm(\mathcal{X})$  and  $E^\pm(\mathcal{X})$  are  $\sigma$ -compact subsets of  $\mathcal{M}$ .*

*Proof.* By assumption,  $\mathcal{X} = \bigcup_{m=1}^{\infty} \mathcal{X}_m$ , where for any  $m \in \mathbb{N}$ ,  $\mathcal{X}_m \subseteq \mathcal{M}$  is closed. Observe that, by (14),

$$\begin{aligned} J^+(\mathcal{X}) &:= \text{pr}_2 \left( \left( \bigcup_{m=1}^{\infty} \mathcal{X}_m \times \mathcal{M} \right) \cap \bigcup_{n=1}^{\infty} \bigcup_{i_1, i_2, \dots, i_n \in \mathbb{N}} J^+_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})} \right) \\ &= \text{pr}_2 \left( \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcup_{i_1, i_2, \dots, i_n \in \mathbb{N}} (\mathcal{X}_m \times \mathcal{M}) \cap J^+_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})} \right) \\ &= \bigcup_{m=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcup_{i_1, i_2, \dots, i_n \in \mathbb{N}} \text{pr}_2 \left( (\mathcal{X}_m \times \mathcal{M}) \cap J^+_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})} \right). \end{aligned}$$

For any  $m, n \in \mathbb{N}$  and any  $i_1, i_2, \dots, i_n \in \mathbb{N}$  the set  $(\mathcal{X}_m \times \mathcal{M}) \cap J^+_{(\overline{N}_{i_1}, \overline{N}_{i_2}, \dots, \overline{N}_{i_n})}$  is closed in  $\overline{N}_{i_1} \times \overline{N}_{i_n}$  and hence compact in  $\mathcal{M}^2$ . Since  $\text{pr}_2$  is a continuous map, the projection of a compact set is itself compact and we obtain that  $J^+(\mathcal{X})$  is  $\sigma$ -compact.

The proof for  $J^-(\mathcal{X})$  is completely analogous. Moreover, on the strength of the previous corollary, replacing  $J^{\pm}$  with  $E^{\pm}$  in the above proof yields the desired result for the horismotical futures and pasts.  $\square$

The final corollary shows that the volume functions can be defined by means of causal futures/pasts instead of the chronological ones.

**Corollary 4.** *Let  $\mathcal{M}$  be a spacetime and  $\eta \in \mathfrak{P}(\mathcal{M})$  be an admissible measure. Then, the volume functions  $t^{\pm}$  associated to  $\eta$  satisfy  $t^{\pm}(p) = \mp \eta(J^{\pm}(p))$  for all  $p \in \mathcal{M}$ . Moreover,  $\eta(E^{\pm}(p)) = 0$  for all  $p \in \mathcal{M}$ .*

*Proof.* By the previous corollary,  $E^{\pm}(p)$  and  $J^{\pm}(p)$  are Borel sets for any  $p \in \mathcal{M}$  and so the expressions  $\eta(E^{\pm}(p))$  and  $\eta(J^{\pm}(p))$  are well-defined. Since it is true that

$$\forall p \in \mathcal{M} \quad I^-(p) \subseteq J^-(p) \subseteq \overline{J^-(p)} = \overline{I^-(p)} = I^-(p) \cup \partial I^-(p),$$

with  $I^-(p) \cap \partial I^-(p) = \emptyset$ , therefore

$$\begin{aligned} t^-(p) &= \eta(\overline{I^-(p)}) \leq \eta(J^-(p)) \leq \eta(\overline{J^-(p)}) \\ &= \eta(\overline{I^-(p)}) = \eta(I^-(p)) + \underbrace{\eta(\partial I^-(p))}_{=0} = t^-(p), \end{aligned}$$

where we have used the second condition in the definition of an admissible measure. Therefore,  $t^-(p) = \eta(J^-(p))$ . The proof for  $t^+$  is analogous.

Moreover, since  $I^{\pm}(p) \subseteq J^{\pm}(p)$  for any  $p \in \mathcal{M}$ ,  $\eta(E^{\pm}(p)) = \eta(J^{\pm}(p) \setminus I^{\pm}(p)) = \eta(J^{\pm}(p)) - \eta(I^{\pm}(p)) = 0$ .  $\square$

## 4. Causality for Probability Measures

The aim of this section is to extend the causal precedence relation  $\preceq$  onto the space of measures  $\mathfrak{P}(\mathcal{M})$  on a given spacetime  $\mathcal{M}$ . We begin by invoking a certain characterisation of causality between events.

Let  $\mathcal{C}(\mathcal{M})$  denote the set of smooth bounded causal functions on the spacetime  $\mathcal{M}$ .

**Theorem 5.** *Let  $\mathcal{M}$  be a globally hyperbolic spacetime. For any  $p, q \in \mathcal{M}$  the following conditions are equivalent*

$$\begin{aligned} 1^\diamond \quad & \forall f \in \mathcal{C}(\mathcal{M}) \quad f(p) \leq f(q), \\ 2^\diamond \quad & p \preceq q. \end{aligned}$$

The proof can be found in [25, Proposition 10].<sup>6</sup>

As an important side note, observe that Theorem 5 exactly mirrors the definition of a causal function. Indeed, the latter can be written symbolically as

$$f \text{ a causal function iff } \forall (p, q) \in J^+ \quad f(p) \leq f(q),$$

whereas Theorem 5 in fact says that

$$(p, q) \in J^+ \text{ iff } \forall f \text{ a causal function } f(p) \leq f(q).$$

Therefore, instead of using  $\preceq$  to define what a causal function is, one can come up with an abstract, suitably structured set  $\mathcal{C}$  of ‘smooth bounded causal functions’ and *define*  $\preceq$  through  $\mathcal{C}$  using the analogue of Theorem 5. This was done by Franco and Eckstein in [25] in a very general context of noncommutative geometry.

Condition  $1^\diamond$  provides a ‘dual’ definition of the causal precedence, which actually suggests how  $\preceq$  could be extended onto  $\mathfrak{P}(\mathcal{M})$ .

**Definition 1.** Let  $\mathcal{M}$  be a globally hyperbolic spacetime. For any  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$  we say that  $\mu$  *causally precedes*  $\nu$  (symbolically  $\mu \preceq \nu$ ) iff

$$\forall f \in \mathcal{C}(\mathcal{M}) \quad \int_{\mathcal{M}} f \, d\mu \leq \int_{\mathcal{M}} f \, d\nu.$$

In [25] it is proven (in a much more general context) that the above-defined relation is in fact a partial order. This definition, however, has two shortcomings. Firstly, it is well motivated only on globally hyperbolic spacetimes. Secondly, the intuitive notion of causality for spread objects, as phrased in the introduction, is not directly visible in Definition 1.

#### 4.1. Characterisations of the Causal Relation

In the following, we provide various conditions which are equivalent to the above definition of a causal relation between measures. Moreover, in some of the implications the assumption on global hyperbolicity of  $\mathcal{M}$  can be relaxed.

The first result states that if  $\mathcal{C}(\mathcal{M})$  is sufficiently rich, one can abandon the smoothness requirement.

**Theorem 6.** *Let  $\mathcal{M}$  be a stably causal spacetime. For any  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$  the following conditions are equivalent:*

<sup>6</sup> The idea that one can recover the causal order from a specific set of functions was developed in [42] (see also [11]). In this light, Theorem 5 can be seen as a simple consequence of [42, Theorem 7]—cf. [41].

1<sup>•</sup> For all  $f \in \mathcal{C}(\mathcal{M})$

$$\int_{\mathcal{M}} f \, d\mu \leq \int_{\mathcal{M}} f \, d\nu. \quad (15)$$

2<sup>•</sup> For all causal  $f \in C_b(\mathcal{M})$

$$\int_{\mathcal{M}} f \, d\mu \leq \int_{\mathcal{M}} f \, d\nu. \quad (16)$$

*Proof.* (1<sup>•</sup>  $\Rightarrow$  2<sup>•</sup>) Relying on [18, Corollary 5.4 and the subsequent comments] we use the fact that in stably causal spacetimes any time function can be uniformly approximated by a *smooth* time (or even temporal) function.

Using the stable causality, fix a temporal function  $\mathcal{T} : \mathcal{M} \rightarrow \mathbb{R}$ . For any  $\varepsilon > 0$ , the function  $f + \varepsilon \arctan \mathcal{T}$  is a time function which clearly approximates  $f$  uniformly. By the above-mentioned corollary, this function in turn can be approximated by a smooth time function  $f_\varepsilon$  such that

$$\forall p \in \mathcal{M} \quad |f(p) + \varepsilon \arctan \mathcal{T}(p) - f_\varepsilon(p)| < \varepsilon. \quad (17)$$

Clearly,  $f_\varepsilon \in \mathcal{C}(\mathcal{M})$ , therefore by 1<sup>•</sup>

$$\int_{\mathcal{M}} f_\varepsilon \, d\mu \leq \int_{\mathcal{M}} f_\varepsilon \, d\nu.$$

To obtain 2<sup>•</sup> it now remains to observe that for any measure  $\eta \in \mathfrak{P}(\mathcal{M})$  it is true that  $\lim_{\varepsilon \rightarrow 0^+} \int_{\mathcal{M}} f_\varepsilon \, d\eta = \int_{\mathcal{M}} f \, d\eta$ .

Indeed, for any  $\eta \in \mathfrak{P}(\mathcal{M})$  and  $\varepsilon > 0$  one has

$$\begin{aligned} \left| \int_{\mathcal{M}} f \, d\eta - \int_{\mathcal{M}} f_\varepsilon \, d\eta \right| &\leq \int_{\mathcal{M}} |f - f_\varepsilon| \, d\eta \\ &\leq \int_{\mathcal{M}} |f + \varepsilon \arctan \mathcal{T} - f_\varepsilon| \, d\eta + \varepsilon \int_{\mathcal{M}} |\arctan \mathcal{T}| \, d\eta \\ &\leq \varepsilon \left( 1 + \frac{\pi}{2} \right), \end{aligned}$$

where we have used (17).

(2<sup>•</sup>  $\Rightarrow$  1<sup>•</sup>) Trivial. □

The next result characterises the relation  $\preceq$  between measures in terms of open future sets.

**Theorem 7.** *Let  $\mathcal{M}$  be a causally continuous spacetime. For any  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$  conditions 1<sup>•</sup> and 2<sup>•</sup> are equivalent to the following condition*

3<sup>•</sup> For every open future set  $\mathcal{F} \subseteq \mathcal{M}$

$$\mu(\mathcal{F}) \leq \nu(\mathcal{F}). \quad (18)$$

*Proof.* ( $2^\bullet \Rightarrow 3^\bullet$ ) Fix an open future set  $\mathcal{F} \subseteq \mathcal{M}$  and let  $\eta$  be an admissible measure on  $\mathcal{M}$ . For any  $\lambda \in (0, 1]$  construct a new admissible measure  $\eta_\lambda := \lambda\eta + (1 - \lambda)\eta(\cdot \cap \mathcal{F})$  and consider the associated past volume function  $t_\lambda^-$  defined via

$$\begin{aligned} \forall p \in \mathcal{M} \quad t_\lambda^-(p) &:= \eta_\lambda(I^-(p)) = \lambda\eta(I^-(p)) + (1 - \lambda)\eta(I^-(p) \cap \mathcal{F}) \\ &= \eta(I^-(p) \cap \mathcal{F}) + \lambda\eta(I^-(p) \setminus \mathcal{F}). \end{aligned}$$

Because  $\mathcal{M}$  is causally continuous,  $t_\lambda^-$  is a time function for any  $\lambda \in (0, 1]$ .

Now, for every  $n \in \mathbb{N}$  define an increasing function  $\varphi_n \in C_b^\infty(\mathbb{R})$  by

$$\forall x \in \mathbb{R} \quad \varphi_n(x) := \frac{1}{2} + \frac{1}{\pi} \arctan(n^2 x - n).$$

The sequence of functions  $(\varphi_n)$  is pointwise convergent to the indicator function of  $\mathbb{R}_{>0}$ . Moreover, also  $\varphi_n \circ t_\lambda^-$  is a bounded time function for every  $n \in \mathbb{N}$  and  $\lambda \in (0, 1]$ . By  $2^\bullet$ , this means that

$$\begin{aligned} \int_{\mathcal{M}} \varphi_n(\eta(I^-(p) \cap \mathcal{F}) + \lambda\eta(I^-(p) \setminus \mathcal{F})) \, d\mu(p) \\ \leq \int_{\mathcal{M}} \varphi_n(\eta(I^-(p) \cap \mathcal{F}) + \lambda\eta(I^-(p) \setminus \mathcal{F})) \, d\nu(p). \end{aligned}$$

Since the functions  $\varphi_n$  are bounded and continuous, we can invoke Lebesgue's dominated convergence theorem and first take  $\lambda \rightarrow 0^+$ , obtaining

$$\int_{\mathcal{M}} \varphi_n(\eta(I^-(p) \cap \mathcal{F})) \, d\mu(p) \leq \int_{\mathcal{M}} \varphi_n(\eta(I^-(p) \cap \mathcal{F})) \, d\nu(p)$$

and then take  $n \rightarrow +\infty$ , which yields

$$\int_{\mathcal{M}} \mathbf{1}_{\mathbb{R}_{>0}}(\eta(I^-(p) \cap \mathcal{F})) \, d\mu(p) \leq \int_{\mathcal{M}} \mathbf{1}_{\mathbb{R}_{>0}}(\eta(I^-(p) \cap \mathcal{F})) \, d\nu(p).$$

It is now crucial to notice that the function  $p \mapsto \eta(\mathcal{F} \cap I^-(p))$  is *positive* on  $\mathcal{F}$  and *zero* on  $\mathcal{M} \setminus \mathcal{F}$ . These observations follow from the definition of an admissible measure and the fact that  $\mathcal{F}$  is future set. Together with the above inequality of integrals they imply that

$$\mu(\mathcal{F}) \leq \nu(\mathcal{F}).$$

( $3^\bullet \Rightarrow 2^\bullet$ ) Let  $f \in C_b(\mathcal{M})$  be causal and let  $\mathcal{T}$  be a temporal function<sup>7</sup> on  $\mathcal{M}$ . For any  $\varepsilon > 0$  define a bounded time function  $f_\varepsilon := f + \varepsilon \arctan \mathcal{T}$ .

Denote  $m := \inf_{p \in \mathcal{M}} f_\varepsilon(p)$  and  $M := \sup_{p \in \mathcal{M}} f_\varepsilon(p)$ . For any fixed  $n \in \mathbb{N}$  define the sets

$$\mathcal{F}_k^{(n)} := f_\varepsilon^{-1}\left(\left(m + k \frac{M-m}{n}, +\infty\right)\right), \quad k = 1, 2, \dots, n-1.$$

<sup>7</sup> Such a function exists because causal continuity implies stable causality. In fact, in the proof of ( $3^\bullet \Rightarrow 2^\bullet$ ) we only need  $\mathcal{M}$  to be stably causal.

Because  $f_\varepsilon$  is continuous and causal, all  $\mathcal{F}_k^{(n)}$ 's are open future sets (cf. Proposition 3).

For any fixed  $n \in \mathbb{N}$  let us consider the following simple function

$$s_n := m + \sum_{k=1}^{n-1} \frac{M-m}{n} \mathbf{1}_{\mathcal{F}_k^{(n)}}.$$

By  $3^\bullet$ , we obtain the following inequality of integrals

$$\int_{\mathcal{M}} s_n \, d\mu = m + \sum_{k=1}^{n-1} \frac{M-m}{n} \mu \left( \mathcal{F}_k^{(n)} \right) \leq m + \sum_{k=1}^{n-1} \frac{M-m}{n} \nu \left( \mathcal{F}_k^{(n)} \right) = \int_{\mathcal{M}} s_n \, d\nu. \quad (19)$$

It is not difficult to realise that

$$\forall p \in \mathcal{M} \quad [\forall n \in \mathbb{N} \quad s_n(p) < f_\varepsilon(p)] \quad \text{and} \quad \lim_{n \rightarrow +\infty} s_n(p) = f_\varepsilon(p).$$

More concretely, one can show that

$$\forall p \in \mathcal{M} \quad f_\varepsilon(p) - s_n(p) \in (0, \frac{M-m}{n}]. \quad (20)$$

Indeed, the very definition of  $\mathcal{F}_k^{(n)}$ 's implies that  $\mathcal{F}_1^{(n)} \supset \mathcal{F}_2^{(n)} \supset \dots \supset \mathcal{F}_{n-1}^{(n)}$ ; therefore, if  $p \in \mathcal{F}_k^{(n)}$  for some  $k \in \{1, \dots, n-1\}$ , then  $p \in \mathcal{F}_j^{(n)}$  for all  $j \in \{1, \dots, k\}$ . This implies that

$$\begin{aligned} s_n(p) &= m + \sum_{k=1}^{n-1} \frac{M-m}{n} \mathbf{1}_{\mathcal{F}_k^{(n)}}(p) = m + \frac{M-m}{n} \max \left\{ k \mid p \in \mathcal{F}_k^{(n)} \right\} \\ &= m + \frac{M-m}{n} \max \left\{ k \mid m + k \frac{M-m}{n} < f_\varepsilon(p) \right\} \\ &= m + \frac{M-m}{n} \max \left\{ k \mid k < \frac{n}{M-m} (f_\varepsilon(p) - m) \right\} \\ &= m + \frac{M-m}{n} \left\lceil \frac{n}{M-m} (f_\varepsilon(p) - m) - 1 \right\rceil, \end{aligned}$$

where  $\lceil \cdot \rceil$  denotes the ceiling function. Using the fact that  $x - \lceil x - 1 \rceil \in (0, 1]$  for any  $x \in \mathbb{R}$ , we obtain that

$$\begin{aligned} f_\varepsilon(p) - s_n(p) &= \frac{M-m}{n} \left( \frac{n}{M-m} (f_\varepsilon(p) - m) - \left\lceil \frac{n}{M-m} (f_\varepsilon(p) - m) - 1 \right\rceil \right) \in (0, \frac{M-m}{n}], \end{aligned}$$

which proves (20).

Invoking now Lebesgue's dominated convergence theorem and passing with  $n \rightarrow +\infty$  in (19) we obtain

$$\int_{\mathcal{M}} f_\varepsilon \, d\mu \leq \int_{\mathcal{M}} f_\varepsilon \, d\nu.$$

Invoking Lebesgue's theorem again, we pass with  $\varepsilon \rightarrow 0^+$  and obtain  $2^\bullet$ .  $\square$

The third and the most important result concerns causally simple spacetimes. We show that condition  $3^\bullet$  extends to different kinds of future sets. Moreover, we introduce a condition that uses the existential quantifier.

**Theorem 8.** *Let  $\mathcal{M}$  be a causally simple spacetime. For any  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$  conditions  $1^\bullet$ ,  $2^\bullet$  and  $3^\bullet$  are equivalent to all of the following conditions*

$4^\bullet$  For every compact  $\mathcal{K} \subseteq \mathcal{M}$

$$\mu(J^+(\mathcal{K})) \leq \nu(J^+(\mathcal{K})). \quad (21)$$

$5^\bullet$  For every Borel future set  $\mathcal{F} \subseteq \mathcal{M}$

$$\mu(\mathcal{F}) \leq \nu(\mathcal{F}). \quad (22)$$

$6^\bullet$  For all  $\varphi, \psi \in C_b(\mathcal{M})$

$$[\forall p, q \in \mathcal{M} \quad p \preceq q \Rightarrow \varphi(p) \leq \psi(q)] \quad \Rightarrow \quad \int_{\mathcal{M}} \varphi \, d\mu \leq \int_{\mathcal{M}} \psi \, d\nu. \quad (23)$$

$7^\bullet$  There exists  $\omega \in \mathfrak{P}(\mathcal{M}^2)$  such that

(i)  $(\text{pr}_1)_*\omega = \mu$  and  $(\text{pr}_2)_*\omega = \nu$ ;

(ii)  $\omega(J^+) = 1$ .

*Proof.* ( $3^\bullet \Rightarrow 4^\bullet$ ) Let  $\mathcal{K}$  be a compact subset of  $\mathcal{M}$ . Fix  $n \in \mathbb{N}$  and cover  $\mathcal{K}$  with open balls of radius  $\frac{1}{n}$ , concretely

$$\mathcal{K} \subseteq \bigcup_{x \in \mathcal{K}} B(x, \tfrac{1}{n}).$$

Hence (see Fig. 2),

$$\forall n \in \mathbb{N} \quad J^+(\mathcal{K}) \subseteq J^+\left(\bigcup_{x \in \mathcal{K}} B(x, \tfrac{1}{n})\right) = \bigcup_{x \in \mathcal{K}} J^+(B(x, \tfrac{1}{n})). \quad (24)$$

We claim that

$$J^+(\mathcal{K}) = \bigcap_{n=1}^{\infty} \bigcup_{x \in \mathcal{K}} J^+(B(x, \tfrac{1}{n})). \quad (25)$$

By (24), it suffices to prove the inclusion ‘ $\supseteq$ ’.

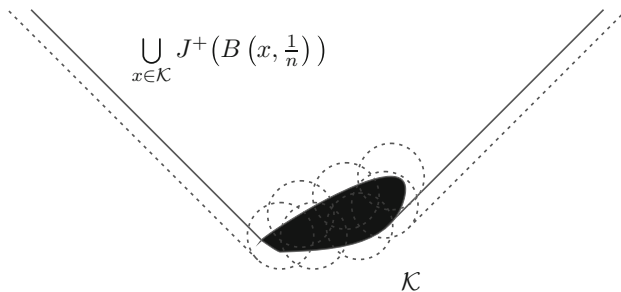


FIGURE 2.  $\{J^+(B(x, \frac{1}{n}))\}_{x \in \mathcal{K}}$  covers  $J^+(\mathcal{K})$

Suppose then that  $q \in \bigcap_{n=1}^{\infty} \bigcup_{x \in \mathcal{K}} J^+ \left( B \left( x, \frac{1}{n} \right) \right)$ , which means that

$$\forall n \in \mathbb{N} \quad \exists x_n \in \mathcal{K} \quad \exists p_n \in B(x_n, \frac{1}{n}) \quad p_n \preceq q.$$

Since  $\mathcal{K}$  is compact, the sequence  $(x_n)$  has a convergent subsequence  $(x_{n_k})$ ,  $\lim_{k \rightarrow +\infty} x_{n_k} = x_{\infty} \in \mathcal{K}$ . Notice that also the subsequence  $(p_{n_k})$  converges to  $x_{\infty}$ . But because  $J^+$  is a closed set in the case of a causally simple spacetime, the fact that for every  $k \in \mathbb{N}$   $p_{n_k} \preceq q$  implies that  $x_{\infty} \preceq q$  and therefore  $q \in J^+(\mathcal{K})$ .

By  $3^\bullet$  we know that

$$\begin{aligned} \mu \left( \bigcup_{x \in \mathcal{K}} J^+ \left( B \left( x, \frac{1}{n} \right) \right) \right) &= \mu \left( \bigcup_{x \in \mathcal{K}} I^+ \left( B \left( x, \frac{1}{n} \right) \right) \right) \\ &\leq \nu \left( \bigcup_{x \in \mathcal{K}} I^+ \left( B \left( x, \frac{1}{n} \right) \right) \right) = \nu \left( \bigcup_{x \in \mathcal{K}} J^+ \left( B \left( x, \frac{1}{n} \right) \right) \right) \end{aligned} \quad (26)$$

Since for all  $n \in \mathbb{N}$ ,  $x \in \mathcal{K}$  we have  $J^+ \left( B \left( x, \frac{1}{n} \right) \right) \supseteq J^+ \left( B \left( x, \frac{1}{n+1} \right) \right)$ , by (2) we obtain that

$$\begin{aligned} \mu(J^+(\mathcal{K})) &= \mu \left( \bigcap_{n=1}^{\infty} \bigcup_{x \in \mathcal{K}} J^+ \left( B \left( x, \frac{1}{n} \right) \right) \right) = \lim_{n \rightarrow +\infty} \mu \left( \bigcup_{x \in \mathcal{K}} J^+ \left( B \left( x, \frac{1}{n} \right) \right) \right) \\ &\leq \lim_{n \rightarrow +\infty} \nu \left( \bigcup_{x \in \mathcal{K}} J^+ \left( B \left( x, \frac{1}{n} \right) \right) \right) = \nu \left( \bigcap_{n=1}^{\infty} \bigcup_{x \in \mathcal{K}} J^+ \left( B \left( x, \frac{1}{n} \right) \right) \right) \\ &= \nu(J^+(\mathcal{K})), \end{aligned}$$

where we have also used (25) and (26), thus proving  $4^\bullet$ .

( $4^\bullet \Rightarrow 5^\bullet$ ) Let  $\mathcal{F} \subseteq \mathcal{M}$  be any Borel future set. For any  $\mathcal{K} \subseteq \mathcal{F}$  it is then true that  $J^+(\mathcal{K}) \subseteq \mathcal{F}$ . Therefore,

$$\mu(\mathcal{K}) \leq \mu(J^+(\mathcal{K})) \leq \mu(\mathcal{F}).$$

In the above chain of inequalities let us take the supremum over all compact  $\mathcal{K} \subseteq \mathcal{F}$ . Using the tightness of  $\mu$  (see (4)), we have

$$\begin{aligned} \mu(\mathcal{F}) &= \sup \{ \mu(\mathcal{K}) \mid \mathcal{K} \subseteq \mathcal{F}, \mathcal{K} \text{ compact} \} \\ &\leq \sup \{ \mu(J^+(\mathcal{K})) \mid \mathcal{K} \subseteq \mathcal{F}, \mathcal{K} \text{ compact} \} \leq \mu(\mathcal{F}), \end{aligned}$$

and so

$$\mu(\mathcal{F}) = \sup \{ \mu(J^+(\mathcal{K})) \mid \mathcal{K} \subseteq \mathcal{F}, \mathcal{K} \text{ compact} \}$$

and similarly for the measure  $\nu$ . As we can see, in order to obtain  $5^\bullet$  from  $4^\bullet$  it is enough to take the supremum over all compact  $\mathcal{K} \subseteq \mathcal{F}$ .

( $5^\bullet \Rightarrow 3^\bullet$ ) Trivial—open sets are Borel.

( $2^\bullet \Rightarrow 6^\bullet$ ) In the first step of the proof we will show that  $6^\bullet$  holds for all nonnegative  $\varphi, \psi \in C_b(\mathcal{M})$  with  $\varphi$  compactly supported. Namely, for such functions we will show that the condition



$$\forall p, q \in \mathcal{M} \quad p \preceq q \Rightarrow \varphi(p) \leq \psi(q) \quad (27)$$

implies the inequality of integrals

$$\int_{\mathcal{M}} \varphi \, d\mu \leq \int_{\mathcal{M}} \psi \, d\nu. \quad (28)$$

Then, in the second step, we will demonstrate that the assumptions of nonnegativity of  $\varphi, \psi$  and of the compactness of  $\text{supp } \varphi$  can in fact be abandoned.

Define a function  $\hat{\varphi} : \mathcal{M} \rightarrow \mathbb{R}$  via  $\hat{\varphi}(p) := \max_{x \preceq p} \varphi(x)$ . Function  $\hat{\varphi}$  is well-defined, because for every  $p \in \mathcal{M}$  the function  $\varphi$ , being continuous, attains its maximum over the compact<sup>8</sup> set  $J^-(p) \cap \text{supp } \varphi$ . Moreover,  $\hat{\varphi}$  satisfies

$$\forall p_1, p_2, q \in \mathcal{M} \quad p_1 \preceq p_2 \preceq q \Rightarrow \varphi(p_1) \leq \hat{\varphi}(p_2) \leq \psi(q). \quad (29)$$

Indeed, the first inequality follows directly from the very definition of  $\hat{\varphi}$ . In order to obtain the second inequality, notice that by (27) we have  $\varphi(p_2) \leq \psi(q)$ . By the transitivity of the relation  $\preceq$ , this inequality holds also if we replace  $p_2$  with any  $x \preceq p_2$ . Hence,

$$\hat{\varphi}(p_2) = \max_{x \preceq p_2} \varphi(x) \leq \psi(q)$$

and (29) is proven.

The function  $\hat{\varphi}$  is obviously nonnegative, bounded, and by the transitivity of  $\preceq$ , it is causal. We claim that it is also continuous.

Indeed, let us show that for any  $\alpha, \beta \in \mathbb{R}$  ( $\alpha < \beta$ ) the preimage  $\hat{\varphi}^{-1}((\alpha, \beta))$  is open.

Notice first that if  $\beta \leq 0$  then, by the nonnegativity of  $\hat{\varphi}$ , the preimage  $\hat{\varphi}^{-1}((\alpha, \beta))$  is empty and hence open. Therefore, we can assume from now on that  $\beta > 0$ .

Observe that  $\hat{\varphi}^{-1}((\alpha, +\infty)) = I^+(\varphi^{-1}((\alpha, +\infty)))$ . This is proven by the following chain of equivalences

$$\begin{aligned} p \in \hat{\varphi}^{-1}((\alpha, +\infty)) &\Leftrightarrow \hat{\varphi}(p) > \alpha \Leftrightarrow \max_{x \preceq p} \varphi(x) > \alpha \\ &\Leftrightarrow \exists x \preceq p \quad \varphi(x) > \alpha \Leftrightarrow \exists x \in \varphi^{-1}((\alpha, +\infty)) \quad x \preceq p \\ &\Leftrightarrow p \in J^+(\varphi^{-1}((\alpha, +\infty))) \end{aligned}$$

and by the observation that, because  $\varphi$  is continuous,  $\varphi^{-1}((\alpha, +\infty))$  is open and hence  $J^+(\varphi^{-1}((\alpha, +\infty))) = I^+(\varphi^{-1}((\alpha, +\infty)))$ .

Similarly, observe that  $\hat{\varphi}^{-1}([\beta, +\infty)) = J^+(\varphi^{-1}([\beta, +\infty)))$ . This is proven by a chain of equivalences analogous to the one above. Notice that because  $\varphi$  is continuous, the preimage  $\varphi^{-1}([\beta, +\infty))$  is closed. Moreover, since  $\varphi$  is nonnegative and  $\beta > 0$ ,  $\varphi^{-1}([\beta, +\infty))$  is contained in the support of  $\varphi$ . But the latter is compact, and so the preimage  $\varphi^{-1}([\beta, +\infty))$ , being a closed subset of a compact set, is itself compact. By the causal simplicity of  $\mathcal{M}$ , the set  $J^+(\varphi^{-1}([\beta, +\infty)))$  is closed.

<sup>8</sup> We are using the fact that in causally simple spacetimes  $J^\pm(p)$  are closed sets for all  $p \in \mathcal{M}$ .

Finally, notice that

$$\begin{aligned}\hat{\varphi}^{-1}((\alpha, \beta)) &= \hat{\varphi}^{-1}((\alpha, +\infty)) \setminus \hat{\varphi}^{-1}([\beta, +\infty)) \\ &= I^+ \left( \varphi^{-1}((\alpha, +\infty)) \right) \setminus J^+ \left( \varphi^{-1}([\beta, +\infty)) \right),\end{aligned}$$

which proves that  $\hat{\varphi}^{-1}((\alpha, \beta))$  is an open set.

We have thus shown that  $\hat{\varphi} \in C_b(\mathcal{M})$ . By  $2^\bullet$  we have that

$$\int_{\mathcal{M}} \hat{\varphi} \, d\mu \leq \int_{\mathcal{M}} \hat{\varphi} \, d\nu. \quad (30)$$

But from (30) we readily obtain (28), because

$$\int_{\mathcal{M}} \varphi \, d\mu \leq \int_{\mathcal{M}} \hat{\varphi} \, d\mu \leq \int_{\mathcal{M}} \hat{\varphi} \, d\nu \leq \int_{\mathcal{M}} \psi \, d\nu,$$

where the first and the last inequalities follow from (29) and the middle one is exactly (30).

Thus, we have already proven  $6^\bullet$  under the assumption that  $\varphi$  is compactly supported and both  $\varphi$  and  $\psi$  are nonnegative. Let us now take any  $\varphi, \psi \in C_b(\mathcal{M})$  satisfying (27).

Define  $m := \min\{\inf \varphi, \inf \psi\}$  and introduce  $\varphi_m, \psi_m \in C_b(\mathcal{M})$  as  $\varphi_m := \varphi - m$  and  $\psi_m := \psi - m$ . Of course,  $\varphi_m, \psi_m \geq 0$ .

Let  $(K_n)_{n \in \mathbb{N}}$  be an exhaustion of  $\mathcal{M}$  by compact sets. Using Urysohn's lemma for LCH spaces (Theorem 1), we construct a sequence  $(\theta_n)_{n \in \mathbb{N}} \subseteq C_c(\mathcal{M})$  of functions such that, for any  $n \in \mathbb{N}$ ,  $\theta_n|_{K_n} \equiv 1$  and  $0 \leq \theta_n \leq 1$ .

Notice that (for every  $n \in \mathbb{N}$ ) the function  $\theta_n \varphi_m$  is compactly supported and, together with  $\psi_m$ , they are nonnegative and satisfy (27), because for all  $p, q \in \mathcal{M}$  such that  $p \preceq q$  one has

$$\theta_n(p) \varphi_m(p) \leq \varphi_m(p) = \varphi(p) - m \leq \psi(q) - m = \psi_m(q).$$

On the strength of the previous part of the proof, it is then true that

$$\int_{\mathcal{M}} \theta_n \varphi_m \, d\mu \leq \int_{\mathcal{M}} \psi_m \, d\nu. \quad (31)$$

By the very definition,  $\theta_n \leq 1$  for every  $n$  and, since  $(K_n)_{n \in \mathbb{N}}$  exhausts  $\mathcal{M}$ , we have that  $\theta_n \rightarrow 1$  pointwise. By Lebesgue's dominated convergence theorem we can pass with  $n \rightarrow +\infty$  in (31) obtaining

$$\int_{\mathcal{M}} \varphi_m \, d\mu \leq \int_{\mathcal{M}} \psi_m \, d\nu.$$

This, in turn, yields

$$\int_{\mathcal{M}} (\varphi(p) - m) \, d\mu(p) \leq \int_{\mathcal{M}} (\psi(q) - m) \, d\nu(q),$$

which, by the fact that  $\mu, \nu$  are probability measures, simplifies to

$$\int_{\mathcal{M}} \varphi \, d\mu \leq \int_{\mathcal{M}} \psi \, d\nu$$

and the proof of  $6^\bullet$  is complete.

( $6^\bullet \Rightarrow 7^\bullet$ ) We will use one of the classical results in the optimal transport theory, concerning what is known as the Kantorovich duality. Concretely, we need the following result adapted from [58, Theorem 1.3].

**Theorem 9** (Kantorovich duality). *Let  $(\mathcal{X}_1, \mu_1)$  and  $(\mathcal{X}_2, \mu_2)$  be two Polish probability spaces and let  $c : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow [0, +\infty]$  be a lower semi-continuous function. Then*

$$\min_{\pi \in \Pi(\mu_1, \mu_2)} \int_{\mathcal{X}_1 \times \mathcal{X}_2} c \, d\pi = \sup_{(\varphi, \psi) \in \Psi(\mu_1, \mu_2)} \left( \int_{\mathcal{X}_1} \varphi \, d\mu_1 - \int_{\mathcal{X}_2} \psi \, d\mu_2 \right), \quad (32)$$

where

- $\Pi(\mu_1, \mu_2) := \{\pi \in \mathfrak{P}(\mathcal{X}_1 \times \mathcal{X}_2) \mid (\text{pr}_i)_* \pi = \mu_i, \, i = 1, 2\},$
- $\Psi(\mu_1, \mu_2) := \{(\varphi, \psi) \in C_b(\mathcal{X}_1) \times C_b(\mathcal{X}_2) \mid \forall x \in \mathcal{X}_1 \, \forall y \in \mathcal{X}_2$   
 $\varphi(x) - \psi(y) \leq c(x, y)\}.$

Let us apply the above theorem to the setting in which  $(\mathcal{X}_1, \mu_1) := (\mathcal{M}, \mu)$ ,  $(\mathcal{X}_2, \mu_2) := (\mathcal{M}, \nu)$  and  $c : \mathcal{M}^2 \rightarrow [0, +\infty]$  is defined as

$$c(p, q) = \begin{cases} 0 & \text{if } p \preceq q \\ +\infty & \text{if } p \not\preceq q \end{cases}.$$

The assumptions of Theorem 9 are met.  $\mathcal{M}$  is a Polish space (cf. Sect. 2.1), whereas the function  $c$  is lower semi-continuous, because the causal simplicity of  $\mathcal{M}$  implies that  $J^+$  is a *closed* subset of  $\mathcal{M}^2$ .

Notice that in the above setting

$$\Psi(\mu, \nu) = \{(\varphi, \psi) \in C_b(\mathcal{M}) \times C_b(\mathcal{M}) \mid \forall p, q \in \mathcal{M} \quad p \preceq q \Rightarrow \varphi(p) \leq \psi(q)\}.$$

In other words,  $\Psi(\mu, \nu)$  is the set of exactly those pairs of functions which satisfy the assumptions of condition  $6^\bullet$ . Since we assume that  $6^\bullet$  holds, we obtain that

$$\forall (\varphi, \psi) \in \Psi(\mu, \nu) \quad \int_{\mathcal{M}} \varphi \, d\mu - \int_{\mathcal{M}} \psi \, d\nu \leq 0$$

and, therefore,

$$\sup_{(\varphi, \psi) \in \Psi(\mu, \nu)} \left( \int_{\mathcal{M}} \varphi \, d\mu - \int_{\mathcal{M}} \psi \, d\nu \right) \leq 0.$$

Using Kantorovich duality (32), we thus obtain that

$$\min_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{M}^2} c(p, q) \, d\pi(p, q) \leq 0.$$

In particular, there exists at least one  $\omega \in \Pi(\mu, \nu)$  such that the integral above is *finite*. But, by the very definition of the function  $c$ , this is possible iff  $\omega(\mathcal{M}^2 \setminus J^+) = 0$  or, equivalently, iff  $\omega(J^+) = 1$ . Thus, we have proven the existence of a measure  $\omega$  with desired properties.

( $7^\bullet \Rightarrow 2^\bullet$ ) Let  $f \in C_b(\mathcal{M})$  be a causal function. Because the probability measures  $\mu$  and  $\nu$  are, respectively, left and right marginals of the joint distribution  $\omega$ , one can write that

$$\begin{aligned} \int_{\mathcal{M}} f(p) d\mu(p) &= \int_{\mathcal{M}^2} f(p) d\omega(p, q) = \int_{J^+} f(p) d\omega(p, q) \\ &\leq \int_{J^+} f(q) d\omega(p, q) = \int_{\mathcal{M}^2} f(q) d\omega(p, q) = \int_{\mathcal{M}} f(q) d\nu(q), \end{aligned}$$

where the inequality follows from the causality of  $f$ . In the integrals with respect to  $\omega$  we can always switch between  $\mathcal{M}^2$  and  $J^+$  because  $\omega(\mathcal{M}^2 \setminus J^+) = 0$ .  $\square$

The fourth result strengthens condition  $5^\bullet$  in the case of globally hyperbolic spacetimes.

**Theorem 10.** *Let  $\mathcal{M}$  be a globally hyperbolic spacetime. Conditions  $1^\bullet$ – $7^\bullet$  are equivalent to both of the following conditions*

$8^\bullet$  *For every Cauchy hypersurface<sup>9</sup>  $\mathcal{S} \subseteq \mathcal{M}$*

$$\mu(J^+(\mathcal{S})) \leq \nu(J^+(\mathcal{S})). \quad (33)$$

$9^\bullet$  *For every compact achronal connected set  $C \subseteq \mathcal{M}$*

$$\mu(J^+(C)) \leq \nu(J^+(C)). \quad (34)$$

*Proof.* ( $5^\bullet \Rightarrow 9^\bullet$ ) Trivial.

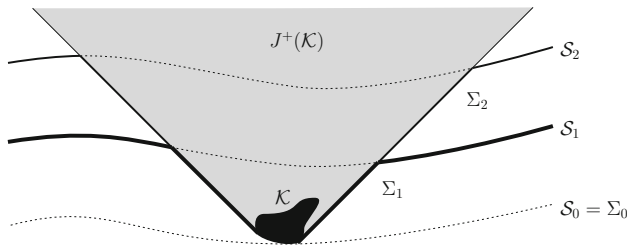
( $9^\bullet \Rightarrow 8^\bullet$ ) Let  $\mathcal{S} \subseteq \mathcal{M}$  be a Cauchy hypersurface. As such,  $\mathcal{S}$  is a connected achronal closed topological hypersurface, thus in particular a locally connected, second-countable LCH space. By Lemma 1,  $\mathcal{S}$  admits an exhaustion  $(C_n)_{n \in \mathbb{N}}$  by compact connected subsets. Of course, each  $C_n$  regarded as a subset of  $\mathcal{M}$  is also compact, connected and achronal. By assumption, we have that

$$\forall n \in \mathbb{N} \quad \mu(J^+(C_n)) \leq \nu(J^+(C_n)).$$

But since  $(C_n)_{n \in \mathbb{N}}$  is increasing, we also have that  $J^+(C_n) \subseteq J^+(C_{n+1})$  for all  $n \in \mathbb{N}$ . Thus, by (1) we obtain that

$$\begin{aligned} \mu(J^+(\mathcal{S})) &= \mu\left(J^+\left(\bigcup_{n=1}^{\infty} C_n\right)\right) = \mu\left(\bigcup_{n=1}^{\infty} J^+(C_n)\right) = \lim_{n \rightarrow +\infty} \mu(J^+(C_n)) \\ &\leq \lim_{n \rightarrow +\infty} \nu(J^+(C_n)) = \nu\left(\bigcup_{n=1}^{\infty} J^+(C_n)\right) = \nu(J^+(\mathcal{S})). \end{aligned}$$

<sup>9</sup> This includes nonsmooth and nonspacelike ones (considered Cauchy surfaces must be achronal, but need not be acausal).

FIGURE 3. The construction of  $\Sigma_n$ 's

( $8^\bullet \Rightarrow 4^\bullet$ ) Let  $\mathcal{T} : \mathcal{M} \rightarrow \mathbb{R}$  be a smooth temporal function, the level sets of which are Cauchy hypersurfaces.

Take any compact subset  $\mathcal{K} \subseteq \mathcal{M}$ . Let  $T_0$  denote the minimal value attained at  $\mathcal{K}$  by the function  $\mathcal{T}$ . For any  $n \in \mathbb{N}_0$  define the level set  $\mathcal{S}_n := \mathcal{T}^{-1}(T_0 + n)$ . Every  $\mathcal{S}_n$  is a smooth spacelike Cauchy hypersurface. Now, for any  $n \in \mathbb{N}_0$  consider the set (see Fig. 3)

$$\Sigma_n := \partial J^+(\mathcal{S}_n \cup \mathcal{K}).$$

We claim that for every  $n \in \mathbb{N}_0$ ,  $\Sigma_n$  is a Cauchy hypersurface and that

$$J^+(\Sigma_n) = J^+(\mathcal{S}_n \cup \mathcal{K}). \quad (35)$$

Indeed, observe first that  $J^+(\mathcal{S}_n \cup \mathcal{K})$  is a future set. By [46, Chapter 14, Corollary 27]  $\Sigma_n$  is therefore a closed achronal topological hypersurface. Let  $\gamma$  be any inextendible timelike curve. It crosses the Cauchy hypersurface  $\mathcal{S}_n$  (which is contained in  $J^+(\mathcal{S}_n \cup \mathcal{K})$ ) and  $\mathcal{S}_0$  (the past of which,  $I^-(\mathcal{S}_0)$ , is disjoint with  $J^+(\mathcal{S}_n \cup \mathcal{K})$ ); therefore, it must cross the boundary  $\partial J^+(\mathcal{S}_n \cup \mathcal{K}) = \Sigma_n$ . Since the latter is achronal, it is met by  $\gamma$  exactly once and therefore  $\Sigma_n$  is a Cauchy hypersurface.

In order to obtain (35), we prove the following lemma.

**Lemma 2.** *Let  $\mathcal{M}$  be a spacetime and let  $\mathcal{F} \subseteq \mathcal{M}$  be a closed future set such that  $\mathcal{F} \subseteq J^+(\mathcal{X})$  for some achronal set  $\mathcal{X}$ . Then,  $J^+(\partial\mathcal{F}) = \mathcal{F}$ .*

*Proof.* ‘ $\subseteq$ ’ Because  $\mathcal{F}$  is closed, it contains its boundary:  $\partial\mathcal{F} \subseteq \mathcal{F}$ . Hence,

$$J^+(\partial\mathcal{F}) \subseteq J^+(\mathcal{F}) = \mathcal{F},$$

because  $\mathcal{F}$  is a future set.

‘ $\supseteq$ ’ Take  $q \in \mathcal{F}$ . By assumption, there exists  $x \in \mathcal{X}$  and a future-directed causal curve  $\gamma$  from  $x$  to  $q$ .

Notice first that  $x \notin \mathcal{F} \setminus \partial\mathcal{F} = \text{int } \mathcal{F}$ . Indeed, if  $x$  would belong to  $\text{int } \mathcal{F}$ , which is an open subset of  $\mathcal{F}$ , there would exist  $x' \in \mathcal{F}$  such that  $x' \ll x$ . But since  $\mathcal{F} \subseteq J^+(\mathcal{X})$ , there would exist  $x'' \in \mathcal{X}$  such that  $x'' \preceq x'$ . Altogether, by (5) we would obtain that  $x'' \ll x$ , in contradiction with the achronality of  $\mathcal{X}$ . Therefore, either  $x \in \partial\mathcal{F}$  or  $x \in \mathcal{M} \setminus \mathcal{F}$ .

If  $x \in \partial\mathcal{F}$ , then  $q \in J^+(\partial\mathcal{F})$  and the proof is complete.

On the other hand, if  $x \in \mathcal{M} \setminus \mathcal{F}$ , then the curve  $\gamma$  must cross  $\partial\mathcal{F}$  at some point  $p$ . Of course,  $p \preceq q$  and hence also in this case  $q \in J^+(\partial\mathcal{F})$ .  $\square$

Notice now that  $J^+(\mathcal{S}_n \cup \mathcal{K}) = J^+(\mathcal{S}_n) \cup J^+(\mathcal{K})$  is in fact a *closed*<sup>10</sup> future set such that  $J^+(\mathcal{S}_n \cup \mathcal{K}) \subseteq J^+(\mathcal{S}_0)$ . On the strength of Lemma 2, we obtain (35).

By 8<sup>•</sup>, because  $\Sigma_n$  is a Cauchy hypersurface for any  $n \in \mathbb{N}_0$ , we can write that

$$\mu(J^+(\Sigma_n)) \leq \nu(J^+(\Sigma_n)). \quad (36)$$

Observe that the sequence  $(J^+(\Sigma_n))_{n \in \mathbb{N}_0}$  is decreasing, because for all  $n \in \mathbb{N}_0$

$$\begin{aligned} J^+(\Sigma_{n+1}) &= J^+(\mathcal{S}_{n+1} \cup \mathcal{K}) = J^+(\mathcal{S}_{n+1}) \cup J^+(\mathcal{K}) \\ &= \mathcal{T}^{-1}([T_0 + n + 1, +\infty)) \cup J^+(\mathcal{K}) \\ &\subseteq \mathcal{T}^{-1}([T_0 + n, +\infty)) \cup J^+(\mathcal{K}) = J^+(\Sigma_n), \end{aligned}$$

where we have used (35) and the very definition of  $\mathcal{S}_n$ 's. Property (2) allows us to pass with  $n \rightarrow +\infty$  in (36) and write that

$$\mu\left(\bigcap_{n=0}^{\infty} J^+(\Sigma_n)\right) \leq \nu\left(\bigcap_{n=0}^{\infty} J^+(\Sigma_n)\right). \quad (37)$$

The countable intersection appearing above can be easily shown to be equal to  $J^+(\mathcal{K})$ . Indeed, one has

$$\begin{aligned} \bigcap_{n=0}^{\infty} J^+(\Sigma_n) &= J^+(\mathcal{K}) \cup \bigcap_{n=0}^{\infty} J^+(\mathcal{S}_n) = J^+(\mathcal{K}) \cup \bigcap_{n=0}^{\infty} \mathcal{T}^{-1}([T_0 + n, +\infty)) \\ &= J^+(\mathcal{K}) \cup \mathcal{T}^{-1}\left(\underbrace{\bigcap_{n=0}^{\infty} [T_0 + n, +\infty)}_{=\emptyset}\right) = J^+(\mathcal{K}). \end{aligned}$$

Therefore, (37) yields (21) and the proof of 4<sup>•</sup> is complete. □

We have thus provided 9 different characterisations of a causal relation between probability measures, which are equivalent if the underlying spacetime is globally hyperbolic. Some of the implications hold under lower causality conditions, as demonstrated in Theorems 6, 7 and 8. Let us now discuss other implications not covered in the proofs.

*Remark 1.* Let us first stress that the formulation of conditions 3<sup>•</sup>, 4<sup>•</sup>, 5<sup>•</sup>, 9<sup>•</sup> using the future of a set is just a matter of convention and one could equally well employ the pasts. Concretely, straightforward application of the time inversion (note that such operation changes the relation  $\preceq$  into the opposite one) shows that these conditions are (in any spacetime  $\mathcal{M}$ ) equivalent to the following ones, respectively:

$$\begin{aligned} 3'^{\bullet} \quad &\text{For every open past set } \mathcal{P} \subseteq \mathcal{M} \\ &\mu(\mathcal{P}) \geq \nu(\mathcal{P}). \end{aligned} \quad (38)$$

<sup>10</sup> For the closedness of  $J^+(\mathcal{S}_n)$ , we refer e.g. to [50, Section 10.2.7]. The closedness of  $J^+(\mathcal{K})$ , on the other hand, follows from the causal simplicity of  $\mathcal{M}$ .

4'• For every compact  $\mathcal{K} \subseteq \mathcal{M}$

$$\mu(J^-(\mathcal{K})) \geq \nu(J^-(\mathcal{K})). \quad (39)$$

5'• For every Borel past set  $\mathcal{P} \subseteq \mathcal{M}$

$$\mu(\mathcal{P}) \geq \nu(\mathcal{P}). \quad (40)$$

9'• For every compact achronal connected set  $C \subseteq \mathcal{M}$

$$\mu(J^-(C)) \geq \nu(J^-(C)). \quad (41)$$

*Remark 2.* Clearly, the proof of implication  $7^\bullet \Rightarrow 2^\bullet$  uses neither the causal simplicity of  $\mathcal{M}$  nor the boundedness of the function  $f$ . In fact, it works for any spacetime and for any  $\mu$ - and  $\nu$ -integrable causal function. We can, therefore, write down the following condition

2'• For every causal  $f \in \mathcal{L}^1(\mathcal{M}, \mu) \cap \mathcal{L}^1(\mathcal{M}, \nu)$ ,

$$\int_{\mathcal{M}} f \, d\mu \leq \int_{\mathcal{M}} f \, d\nu. \quad (42)$$

For any spacetime  $\mathcal{M}$  it is then true that  $7^\bullet \Rightarrow 2'^\bullet$  as well as, trivially,  $2'^\bullet \Rightarrow 2^\bullet \Rightarrow 1^\bullet$ .

*Remark 3.* Condition  $2'^\bullet$  implies  $5^\bullet$  in any spacetime  $\mathcal{M}$ .

*Proof.* Let  $\mathcal{F}$  be a Borel future subset of  $\mathcal{M}$ . Clearly,  $1_{\mathcal{F}} \in \mathcal{L}^1(\mathcal{M}, \mu) \cap \mathcal{L}^1(\mathcal{M}, \nu)$  and, by Corollary 1, it is a causal function. By condition  $2'^\bullet$  we can write

$$\mu(\mathcal{F}) = \int_{\mathcal{M}} 1_{\mathcal{F}} \, d\mu \leq \int_{\mathcal{M}} 1_{\mathcal{F}} \, d\nu = \nu(\mathcal{F}),$$

what proves  $5^\bullet$ . □

*Remark 4.* Also implication  $7^\bullet \Rightarrow 6^\bullet$  holds in all spacetimes. We can show even slightly more, namely, that condition  $7^\bullet$  implies

6'• For all  $\varphi, \psi : \mathcal{M} \rightarrow \mathbb{R}$  such that  $\varphi$  is  $\mu$ -integrable and  $\psi$  is  $\nu$ -integrable

$$[\forall p, q \in \mathcal{M} \quad p \preceq q \Rightarrow \varphi(p) \leq \psi(q)] \Rightarrow \int_{\mathcal{M}} \varphi \, d\mu \leq \int_{\mathcal{M}} \psi \, d\nu. \quad (43)$$

*Proof.* Similarly as in the proof of  $7^\bullet \Rightarrow 2^\bullet$ , one can write that

$$\begin{aligned} \int_{\mathcal{M}} \varphi(p) \, d\mu(p) &= \int_{\mathcal{M}^2} \varphi(p) \, d\omega(p, q) = \int_{J^+} \varphi(p) \, d\omega(p, q) \\ &\leq \int_{J^+} \psi(q) \, d\omega(p, q) = \int_{\mathcal{M}^2} \psi(q) \, d\omega(p, q) = \int_{\mathcal{M}} \psi(q) \, d\nu(q), \end{aligned}$$

where the inequality follows from the assumptions on  $\varphi$  and  $\psi$ . □

## 4.2. Basic Properties of the Causal Relation Between Measures

In the previous subsection we have shown that for *any* spacetime  $\mathcal{M}$  the condition  $7^\bullet$  not only implies all of the others listed in Theorems 6, 7, 8 and 10, but also more general ones  $2'^\bullet$  and  $6'^\bullet$ . It encourages us to promote the condition  $7^\bullet$  to the *definition* of the causal precedence relation on  $\mathfrak{P}(\mathcal{M})$  for any spacetime  $\mathcal{M}$ .

**Definition 2.** Let  $\mathcal{M}$  be a spacetime. For any  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$  we say that  $\mu$  *causally precedes*  $\nu$  (symbolically  $\mu \preceq \nu$ ) iff there exists  $\omega \in \mathfrak{P}(\mathcal{M}^2)$  such that

- (i)  $(\text{pr}_1)_*\omega = \mu$  and  $(\text{pr}_2)_*\omega = \nu$ ,
- (ii)  $\omega(J^+) = 1$ .

Such an  $\omega$  will be called a *causal coupling* of  $\mu$  and  $\nu$ .

Observe that  $\omega(J^+)$  is well-defined because, by Theorem 4,  $J^+$  is  $\sigma$ -compact, and hence Borel, for any spacetime  $\mathcal{M}$ .

*Remark 5.* In the case of causally simple spacetimes  $J^+ \subseteq \mathcal{M}^2$  is closed and therefore, by the very definition of the support of a measure (see the last paragraph of Sect. 2.2), condition (ii) in Definition 2 is equivalent to the inclusion  $\text{supp } \omega \subseteq J^+$ . However, without the assumption of causal simplicity this is no longer true.

The term ‘coupling (of measures  $\mu$  and  $\nu$ )’ comes from the optimal transport theory [58], where it describes any  $\omega \in \mathfrak{P}(\mathcal{M}^2)$  with property (i) of the above definition. The set of such couplings, denoted  $\Pi(\mu, \nu)$ , has already appeared above in the context of the Kantorovich duality (Theorem 9).

Such a coupling—or a *transference plan*, as it is also called—can be regarded as an instruction how to ‘reconfigure’ a fixed amount of ‘mass’ distributed over  $\mathcal{M}$  according to the measure  $\mu$  so that it becomes distributed according to the measure  $\nu$ . This ‘reconfiguration’ involves transporting the (possibly infinitesimal) portions of ‘mass’ between points of  $\mathcal{M}$ , and the coupling  $\omega \in \Pi(\mu, \nu) \subseteq \mathfrak{P}(\mathcal{M}^2)$  precisely describes what amount of ‘mass’ is transported between any given pair of points.

It is, however, property (ii) which ties the above definition with causality theory. It can be summarised as a requirement that the transport of ‘mass’ should be conducted along future-directed causal curves only—that is why such couplings deserve to be called *causal*. The set of all causal couplings of measures  $\mu$  and  $\nu$  will be denoted by  $\Pi_c(\mu, \nu)$ .

Notice that a (causal) coupling does *not* specify along *which* (causal) curves the portions of ‘mass’ are transported. In fact, various families of (causal) curves can lead to the same (causal) coupling. Notice also that the ‘mass’ concentrated initially at some point  $p \in \mathcal{M}$  can dilute to many different points.

Observe that for Dirac measures  $\mu = \delta_p$ ,  $\nu = \delta_q$  Definition 2 reduces to the standard definition of the causal relation between events  $p$  and  $q$ . This can be seen as a corollary of the following proposition.



**Proposition 4.** Let  $\mathcal{M}$  be a topological space and let  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$  and  $\omega \in \Pi(\mu, \nu)$ . Then, for any Borel sets  $A, B \subseteq \mathcal{M}$

- (i)  $\mu(A) = \nu(B) = 1 \Leftrightarrow \omega(A \times B) = 1$ ,
- (ii)  $\mu(A) = 0 \vee \nu(B) = 0 \Rightarrow \omega(A \times B) = 0$ .

*Proof.* (i) To prove ‘ $\Rightarrow$ ’ we use the inclusion–exclusion principle to write

$$\begin{aligned} 1 &\geq \omega(A \times B) = \omega(A \times \mathcal{M} \cap \mathcal{M} \times B) \\ &= \underbrace{\omega(A \times \mathcal{M})}_{=\mu(A)=1} + \underbrace{\omega(\mathcal{M} \times B)}_{=\nu(B)=1} - \underbrace{\omega(A \times \mathcal{M} \cup \mathcal{M} \times B)}_{\leq 1} \geq 1 + 1 - 1 = 1. \end{aligned}$$

Conversely, to prove ‘ $\Leftarrow$ ’, notice that

$$\begin{aligned} 1 &\geq \mu(A) = \omega(A \times \mathcal{M}) \geq \omega(A \times B) = 1 \\ \text{and } 1 &\geq \nu(B) = \omega(\mathcal{M} \times B) \geq \omega(A \times B) = 1. \end{aligned}$$

(ii) One has

$$0 \leq \omega(A \times B) \leq \min \{ \omega(A \times \mathcal{M}), \omega(\mathcal{M} \times B) \} = \min \{ \mu(A), \nu(B) \} = 0.$$

□

**Corollary 5.** Let  $\mathcal{M}$  be a spacetime. Then, for any  $p, q \in \mathcal{M}$   $p \preceq q$  iff  $\delta_p \preceq \delta_q$ .

*Proof.* By Proposition 4, the only coupling between two Dirac measures  $\delta_p, \delta_q$  is their product measure  $\omega := \delta_p \times \delta_q = \delta_{(p,q)}$ . Hence, the fact that  $p \preceq q$  is equivalent in this case to the requirement that  $\omega(J^+) = 1$ . □

**Corollary 6.** Let  $\mathcal{M}$  be a causally simple spacetime. For any  $p, q \in \mathcal{M}$  the following conditions are equivalent

- 1 $^\diamond$   $\forall f \in \mathcal{C}(\mathcal{M}) \quad f(p) \leq f(q)$ ,
- 2 $^\diamond$   $p \preceq q$ .

*Proof.* It is a direct consequence of the equivalence ( $1^\bullet \Leftrightarrow 7^\bullet$ ) in Theorem 8 and Corollary 5. □

If the measure  $\mu$  is compactly supported, then in the light of the above discussion it is natural to expect that the support of any  $\nu$  with  $\mu \preceq \nu$  should be within the future of  $\text{supp } \mu$  [60]. This intuitive condition is in fact true in causally simple spacetimes.

**Proposition 5.** Let  $\mathcal{M}$  be a spacetime and let  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$ , with  $\mu$  compactly supported and  $\mu \preceq \nu$ . Then,  $\nu(J^+(\text{supp } \mu)) = 1$ . Moreover, if  $\mathcal{M}$  is causally simple then  $\text{supp } \nu \subseteq J^+(\text{supp } \mu)$ .

*Proof.* By condition 4 $^\bullet$  (which is implied by Definition 2) it is true that

$$1 = \mu(\text{supp } \mu) \leq \mu(J^+(\text{supp } \mu)) \leq \nu(J^+(\text{supp } \mu)) \leq 1$$

and therefore  $\nu(J^+(\text{supp } \mu)) = 1$ .

We now claim that if  $\mathcal{M}$  is causally simple, then this implies that  $\text{supp } \nu \subseteq J^+(\text{supp } \mu)$ .

Indeed, recall that in a causally simple spacetime the causal futures of compact sets are closed. Therefore, if there existed  $q \in \text{supp } \nu$  but  $q \notin$

$J^+(\text{supp } \mu)$ , then we could take an open neighbourhood  $U \ni q$  such that  $\nu(U) > 0$  but  $U \cap J^+(\text{supp } \mu) = \emptyset$ . But this would imply that

$$\nu(J^+(\text{supp } \mu)) \leq 1 - \nu(U) < 1,$$

in contradiction with the first part of the proof.  $\square$

Recall that the causal precedence relation between events is reflexive, transitive and, iff  $\mathcal{M}$  is causal, antisymmetric. We now prove analogous results for the space of Borel probability measures on  $\mathcal{M}$  equipped with the relation  $\preceq$ . To this end, it will be convenient to use the *diagonal map*  $\Delta : \mathcal{M} \rightarrow \mathcal{M}^2$ , defined as  $\Delta(p) := (p, p)$  for any  $p \in \mathcal{M}$ .

**Theorem 11.** *Let  $\mathcal{M}$  be a spacetime. The relation  $\preceq$  on  $\mathfrak{P}(\mathcal{M})$  is reflexive and transitive.*

*Proof.* To prove reflexivity of  $\preceq$ , it suffices to notice that for any  $\mu \in \mathfrak{P}(\mathcal{M})$  the pushforward measure  $\Delta_*\mu$  is a causal coupling of  $\mu$  with itself.

Indeed,  $(\text{pr}_i)_*\Delta_*\mu = (\text{pr}_i \circ \Delta)_*\mu = \mu$  for  $i = 1, 2$  and  $\Delta_*\mu(J^+) = \mu(\Delta^{-1}(J^+)) = \mu(\mathcal{M}) = 1$ , where we have used the equality  $\Delta^{-1}(J^+) = \mathcal{M}$ , which expresses nothing but the reflexivity of the causal precedence relation between *events*.

We now move to proving the transitivity of  $\preceq$ . Let us invoke the following standard result [58, Lemma 7.6] from the optimal transport theory.

**Lemma 3** (Gluing Lemma). *Let  $(\mathcal{X}_i, \mu_i)$ ,  $i = 1, 2, 3$  be Polish probability spaces and assume there exist couplings  $\omega_{12} \in \Pi(\mu_1, \mu_2)$  and  $\omega_{23} \in \Pi(\mu_2, \mu_3)$ .*

*Then, there exists  $\omega_{123} \in \mathfrak{P}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3)$  such that  $(\text{pr}_{12})_*\omega_{123} = \omega_{12}$  and  $(\text{pr}_{23})_*\omega_{123} = \omega_{23}$ , where  $\text{pr}_{ij} : \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \rightarrow \mathcal{X}_i \times \mathcal{X}_j$  denotes the canonical projection map.*

*Moreover,  $\omega_{13} := (\text{pr}_{13})_*\omega_{123}$  belongs to  $\Pi(\mu_1, \mu_3)$ .*

The Gluing Lemma works well with the causal precedence relation. Concretely, let us take  $\mu_1, \mu_2, \mu_3 \in \mathfrak{P}(\mathcal{M})$  such that  $\mu_1 \preceq \mu_2 \preceq \mu_3$ , where  $\omega_{12} \in \Pi_c(\mu_1, \mu_2)$  and  $\omega_{23} \in \Pi_c(\mu_2, \mu_3)$ . Then, the coupling  $\omega_{13}$  of  $\mu_1$  and  $\mu_3$  is causal too.

Indeed, notice first that

$$\begin{aligned} \omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \not\preceq r\}) &\leq \omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid q \not\preceq r\}) \\ &= \omega_{123}(\mathcal{M} \times (\mathcal{M}^2 \setminus J^+)) = \omega_{23}(\mathcal{M}^2 \setminus J^+) = 1 - \omega_{23}(J^+) = 0 \end{aligned}$$

and thus  $\omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \not\preceq r\}) = 0$ .

On the other hand,

$$\begin{aligned} \omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid p \not\preceq q\}) &= \omega_{123}((\mathcal{M}^2 \setminus J^+) \times \mathcal{M}) \\ &= \omega_{12}(\mathcal{M}^2 \setminus J^+) = 1 - \omega_{12}(J^+) = 0. \end{aligned}$$

Since  $\mathcal{M}^3$  can be decomposed into the following union of (pairwise disjoint) sets

$$\begin{aligned}\mathcal{M}^3 = & \{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \preceq r\} \cup \{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \not\preceq r\} \\ & \cup \{(p, q, r) \in \mathcal{M}^3 \mid p \not\preceq q\},\end{aligned}$$

we obtain

$$\begin{aligned}1 = \omega_{123}(\mathcal{M}^3) &= \omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \preceq r\}) \\ &+ \underbrace{\omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \not\preceq r\})}_{=0} + \underbrace{\omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid p \not\preceq q\})}_{=0}\end{aligned}$$

and hence

$$\omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \preceq r\}) = 1. \quad (44)$$

But this, in turn, means that

$$\begin{aligned}1 \geq \omega_{13}(J^+) &= \omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq r\}) \\ &\geq \omega_{123}(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \preceq r\}) = 1,\end{aligned}$$

where the middle inequality is a direct consequence of the transitivity of the causal precedence relation between events. We have thus proven that  $\omega_{13}(J^+) = 1$ , and so  $\omega_{13} \in \Pi_c(\mu_1, \mu_3)$  and therefore  $\mu_1 \preceq \mu_3$ .  $\square$

The natural question arises: How robust the causal structure of a space-time  $\mathcal{M}$  must be to render the relation  $\preceq$  antisymmetric and hence a partial order? Obviously,  $\mathcal{M}$  must be at least causal (otherwise even the causal precedence relation between *events* fails to be antisymmetric). We have the following result:

**Theorem 12.** *Let  $\mathcal{M}$  be a spacetime with the following property:*

*For any compact  $\mathcal{K} \subseteq \mathcal{M}$  there exists a Borel function  $\tau_{\mathcal{K}} : \mathcal{K} \rightarrow \mathbb{R}$  such that*

$$\forall p, q \in \mathcal{K} \quad (p \preceq q \wedge p \neq q) \Rightarrow \tau_{\mathcal{K}}(p) < \tau_{\mathcal{K}}(q). \quad (45)$$

*Then, for any  $\mu \in \mathfrak{P}(\mathcal{M})$ ,  $\Pi_c(\mu, \mu) = \{\Delta_*\mu\}$ . Moreover, the relation  $\preceq$  is antisymmetric.*

*Remark 6.* Property (45) implies that  $\mathcal{M}$  is causal. Indeed, suppose that there exist two distinct events  $p, q \in \mathcal{M}$  such that  $p \preceq q \preceq p$ . Now, taking now  $\mathcal{K} = \{p, q\}$  we would obtain, on the strength of (45), that  $\tau_{\mathcal{K}}(p) < \tau_{\mathcal{K}}(q) < \tau_{\mathcal{K}}(p)$ , a contradiction.

On the other hand, if  $\mathcal{M}$  is past (future) distinguishing, then any past (resp. future) volume function is a semi-continuous, and hence Borel, generalised time function  $\tau$  (cf. Sect. 2.3). This obviously implies (45)—for any compact  $\mathcal{K} \subseteq \mathcal{M}$  simply define  $\tau_{\mathcal{K}} := \tau|_{\mathcal{K}}$ . However, being past or future distinguishing is not necessary for (45) to hold.

Indeed, the rightmost diagram in [43, Figure 6], which we reproduce<sup>11</sup> in Fig. 4 for reader's convenience, presents a causal, but neither future nor

<sup>11</sup> We thank Ettore Minguzzi for providing us the diagram from [43].

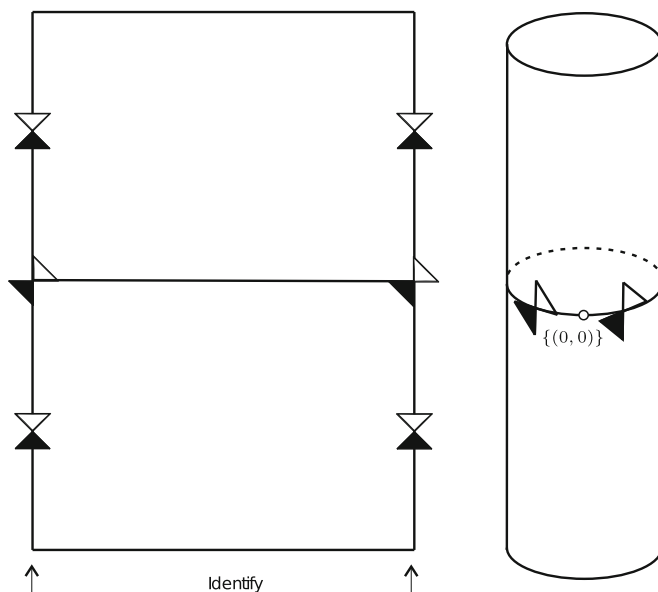


FIGURE 4. [43, Figure 6] An example of a causal, but neither future nor past distinguishing spacetime, which nevertheless admits a Borel generalised time function

past distinguishing spacetime  $\mathcal{M} := \mathbb{R} \times S^1 \setminus \{(0, 0)\}$ , which admits a Borel generalised time function. For instance,

$$\tau(x, \theta) := \begin{cases} \arctan x & \text{for } x < 0 \\ \theta & \text{for } x = 0 \\ 2\pi + \arctan x & \text{for } x > 0 \end{cases},$$

for any  $x \in \mathbb{R}$  and  $\theta \in S^1$ , where the latter is the angular coordinate, the range of which is  $[0, 2\pi)$ , except for  $x = 0$ , when its range is  $(0, 2\pi)$ .

Before we move to the proof of Theorem 12, let us present the following lemma.

**Lemma 4.** *Let  $\mathcal{M}$  be a topological space and let  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$ . Moreover, let  $\omega \in \Pi(\mu, \nu)$  be such that  $\omega(\Delta(\mathcal{M})) = 1$ . Then,  $\mu = \nu$  and  $\omega = \Delta_*\mu = \Delta_*\nu$ .*

*Proof.* Let  $\mathcal{U}$  be any Borel subset of  $\mathcal{M}^2$ . Then,  $\omega(\mathcal{U} \setminus \Delta(\mathcal{M})) \leq \omega(\mathcal{M} \setminus \Delta(\mathcal{M})) = 1 - \omega(\Delta(\mathcal{M})) = 0$  and therefore  $\omega(\mathcal{U} \setminus \Delta(\mathcal{M})) = 0$ . But this allows us to write

$$\omega(\mathcal{U}) = \omega(\mathcal{U} \cap \Delta(\mathcal{M})) + \underbrace{\omega(\mathcal{U} \setminus \Delta(\mathcal{M}))}_{=0} = \omega(\Delta(\Delta^{-1}(\mathcal{U}))).$$

The rightmost expression, in turn, can be further transformed either into

$$\begin{aligned}\omega(\Delta(\Delta^{-1}(\mathcal{U}))) &= \omega((\Delta^{-1}(\mathcal{U}) \times \mathcal{M}) \cap \Delta(\mathcal{M})) \\ &= \omega(\Delta^{-1}(\mathcal{U}) \times \mathcal{M}) - \underbrace{\omega((\Delta^{-1}(\mathcal{U}) \times \mathcal{M}) \setminus \Delta(\mathcal{M}))}_{=0} \\ &= \mu(\Delta^{-1}(\mathcal{U})) = \Delta_*\mu(\mathcal{U})\end{aligned}$$

or into

$$\begin{aligned}\omega(\Delta(\Delta^{-1}(\mathcal{U}))) &= \omega((\mathcal{M} \times \Delta^{-1}(\mathcal{U})) \cap \Delta(\mathcal{M})) \\ &= \omega(\mathcal{M} \times \Delta^{-1}(\mathcal{U})) - \underbrace{\omega(\mathcal{M} \times (\Delta^{-1}(\mathcal{U}) \setminus \Delta(\mathcal{M})))}_{=0} \\ &= \nu(\Delta^{-1}(\mathcal{U})) = \Delta_*\nu(\mathcal{U}),\end{aligned}$$

what proves the second part of the theorem. To obtain the equality  $\mu = \nu$ , take any Borel  $\mathcal{V} \subseteq \mathcal{M}$  and notice, for instance, that

$$\nu(\mathcal{V}) = \omega(\mathcal{M} \times \mathcal{V}) = \Delta_*\mu(\mathcal{M} \times \mathcal{V}) = \mu(\Delta^{-1}(\mathcal{M} \times \mathcal{V})) = \mu(\mathcal{V}),$$

which concludes the entire proof.  $\square$

*Proof of Theorem 12.* Take any  $\mu \in \mathfrak{P}(\mathcal{M})$  and let  $\pi \in \Pi_c(\mu, \mu)$ . By Definition 2, we have that

$$\forall f \in \mathcal{L}^1(\mathcal{M}, \mu) \quad \int_{J^+} f(p) d\pi(p, q) = \int_{\mathcal{M}} f d\mu = \int_{J^+} f(q) d\pi(p, q)$$

and hence

$$\forall f \in \mathcal{L}^1(\mathcal{M}, \mu) \quad \int_{J^+} (f(q) - f(p)) d\pi(p, q) = 0$$

or, by noticing that the integrand vanishes on  $\Delta(\mathcal{M})$ ,

$$\forall f \in \mathcal{L}^1(\mathcal{M}, \mu) \quad \int_{J^+ \setminus \Delta(\mathcal{M})} (f(q) - f(p)) d\pi(p, q) = 0. \quad (46)$$

Suppose now that  $\pi(J^+ \setminus \Delta(\mathcal{M})) > 0$ . Because  $\pi$  is tight, there exists a compact set  $K \subseteq J^+ \setminus \Delta(\mathcal{M})$  with  $\pi(K) > 0$ . Notice that  $K \subseteq \mathcal{K}^2$ , where  $\mathcal{K} := \text{pr}_1 K \cup \text{pr}_2 K$  is a compact subset of  $\mathcal{M}$ , and so  $\pi(\mathcal{K}^2 \cap J^+ \setminus \Delta(\mathcal{M})) > 0$ . Define  $f_{\mathcal{K}} : \mathcal{M} \rightarrow \mathbb{R}$  via

$$f_{\mathcal{K}}(p) := \begin{cases} \arctan \tau_{\mathcal{K}}(p) & \text{for } p \in \mathcal{K} \\ 0 & \text{for } p \notin \mathcal{K} \end{cases},$$

where  $\tau_{\mathcal{K}}$  is a function, which exists by property (45). Function  $f_{\mathcal{K}}$  is Borel and bounded, and hence  $\mu$ -integrable. Plugging it into (46) yields

$$\int_{\mathcal{K}^2 \cap J^+ \setminus \Delta(\mathcal{M})} (\arctan \tau_{\mathcal{K}}(q) - \arctan \tau_{\mathcal{K}}(p)) d\pi(p, q) = 0.$$

But the integrand of the above integral is positive on  $\mathcal{K}^2 \cap J^+ \setminus \Delta(\mathcal{M})$  by the very definition of  $\tau_{\mathcal{K}}$ , therefore the fact that the integral is zero implies

that  $\pi(\mathcal{K}^2 \cap J^+ \setminus \Delta(\mathcal{M})) = 0$ , which contradicts the earlier result. This proves that  $\pi(J^+ \setminus \Delta(\mathcal{M})) = 0$ .

By property (ii) from Definition 2, this in turn means that

$$\pi(\Delta(\mathcal{M})) = \pi(J^+) - \pi(J^+ \setminus \Delta(\mathcal{M})) = 1.$$

On the strength of Lemma 4, we get that  $\pi = \Delta_*\mu$ .

We now move to proving the antisymmetry of the relation  $\preceq$ . Let  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$  be such that  $\mu \preceq \nu \preceq \mu$ . Let  $\omega \in \Pi_c(\mu, \nu)$  and  $\varpi \in \Pi_c(\nu, \mu)$ . By the Gluing Lemma, there exists  $\Omega \in \mathfrak{P}(\mathcal{M}^3)$  such that  $(\text{pr}_{12})_*\Omega = \omega$ ,  $(\text{pr}_{23})_*\Omega = \varpi$  and  $(\text{pr}_{13})_*\Omega \in \Pi_c(\mu, \mu)$ , which, by the previous part of the proof, means that  $(\text{pr}_{13})_*\Omega = \Delta_*\mu$ .

Formula (44) takes here the following form

$$\Omega(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \preceq r\}) = 1.$$

Notice, however, that the set  $\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \preceq r \neq p\}$  is  $\Omega$ -null, because

$$\begin{aligned} \Omega(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \preceq r \neq p\}) &\leq \Omega(\{(p, q, r) \in \mathcal{M}^3 \mid p \neq r\}) \\ &= (\text{pr}_{13})_*\Omega(\{(p, r) \in \mathcal{M}^2 \mid p \neq r\}) = \Delta_*\mu(\mathcal{M}^2 \setminus \Delta(\mathcal{M})) = 1 - \mu(\mathcal{M}) = 0. \end{aligned}$$

Therefore, in fact,

$$\begin{aligned} \Omega(\{(p, q, p) \in \mathcal{M}^3 \mid p \preceq q \preceq p\}) \\ = \underbrace{\Omega(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \preceq r\})}_{=1} - \underbrace{\Omega(\{(p, q, r) \in \mathcal{M}^3 \mid p \preceq q \preceq r \neq p\})}_{=0}. \end{aligned} \quad (47)$$

But  $\mathcal{M}$  is causal (cf. Remark 6), therefore the causal precedence relation between events is antisymmetric, and thus the set, the measure of which is evaluated in (47), is equal to  $\{(p, p, p) \in \mathcal{M}^3 \mid p \in \mathcal{M}\}$ .

We can now easily obtain that

$$\begin{aligned} \omega(\Delta(\mathcal{M})) &= \Omega(\Delta(\mathcal{M}) \times \mathcal{M}) = \Omega(\{(p, p, q) \in \mathcal{M}^3 \mid p, q \in \mathcal{M}\}) \\ &\geq \Omega(\{(p, p, p) \in \mathcal{M}^3 \mid p \in \mathcal{M}\}) = 1 \end{aligned}$$

and so  $\omega(\Delta(\mathcal{M})) = 1$ . Invoking Lemma 4, we obtain that  $\mu = \nu$ .  $\square$

## 5. Lorentz–Wasserstein Distances

Recall that the Lorentzian distance  $d : \mathcal{M}^2 \rightarrow [0, +\infty]$  provides a way of measuring distances between events, in an analogy with the Riemannian distance  $d_R$  in the case of Riemannian manifolds. In the latter case, one can extend the notion of a distance to the space of measures on  $\mathcal{M}$ . Concretely, for any  $s \geq 1$  one defines the so-called  $s^{\text{th}}$  *Wasserstein distance* between any two measures  $\mu, \nu \in \mathfrak{P}(\mathbb{R})$  on a Riemannian manifold  $\mathbb{R}$  as

$$W_s(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \left[ \int_{\mathbb{R}^2} d_R(x, y)^s d\pi(x, y) \right]^{1/s}. \quad (48)$$

For an exposition of the theory of Wasserstein distances in the context of the optimal transport theory one is referred e.g. to [58].

We now propose the following natural definition of a distance between measures on a spacetime.

**Definition 3.** Let  $\mathcal{M}$  be a spacetime and let  $s \in (0, 1]$ . The  $s^{\text{th}}$  Lorentz–Wasserstein distance is the map  $LW_s : \mathfrak{P}(\mathcal{M}) \times \mathfrak{P}(\mathcal{M}) \rightarrow [0, +\infty]$  given by

$$LW_s(\mu, \nu) := \begin{cases} \sup_{\omega \in \Pi_c(\mu, \nu)} \left[ \int_{\mathcal{M}^2} d(p, q)^s d\omega(p, q) \right]^{1/s} & \text{if } \Pi_c(\mu, \nu) \neq \emptyset \\ 0 & \text{if } \Pi_c(\mu, \nu) = \emptyset \end{cases}.$$

Notice that the integrals are well-defined, because  $d$  is lower semi-continuous and hence Borel. Notice also that for Dirac measures  $LW_s(\delta_p, \delta_q) = d(p, q)$  for any  $s$ .

Lorentz–Wasserstein distances have properties analogous to those of the Lorentzian distance (cf. Sect. 2.3).

**Theorem 13.** Let  $\mathcal{M}$  be a spacetime and let  $s \in (0, 1]$ . Then:

(i) For any  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$

$$LW_s(\mu, \nu) > 0 \quad \Leftrightarrow \quad \exists \omega \in \Pi_c(\mu, \nu) \quad \omega(I^+) > 0 \quad \Rightarrow \quad \mu \preceq \nu.$$

(ii) The reverse triangle inequality holds. Namely, for all  $\mu_1, \mu_2, \mu_3 \in \mathfrak{P}(\mathcal{M})$

$$\mu_1 \preceq \mu_2 \preceq \mu_3 \quad \Rightarrow \quad LW_s(\mu_1, \mu_2) + LW_s(\mu_2, \mu_3) \leq LW_s(\mu_1, \mu_3). \quad (49)$$

(iii) For any  $\mu \in \mathfrak{P}(\mathcal{M})$ ,  $LW_s(\mu, \mu)$  is either 0 or  $+\infty$ .

(iv)  $\mathcal{M}$  is chronological iff  $\forall \mu \in \mathfrak{P}(\mathcal{M}) \quad LW_s(\mu, \mu) = 0$ .

(v) For any  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$ , if  $LW_s(\mu, \nu) \in (0, +\infty)$  then  $LW_s(\nu, \mu) = 0$ .

*Proof.* (i) The implication is obvious, so we only prove the equivalence.

To prove the ‘ $\Rightarrow$ ’ part of the equivalence, assume that  $LW_s(\mu, \nu) > 0$ . By the very definition of  $LW_s$ , this implies that there exists  $\omega \in \Pi_c(\mu, \nu)$  such that  $\int_{\mathcal{M}^2} d(p, q)^s d\omega(p, q) > 0$ . In order to prove that  $\omega(I^+) > 0$ , suppose on the contrary that  $I^+$  is  $\omega$ -null. Then, one would have

$$\begin{aligned} 0 &< \int_{\mathcal{M}^2} d(p, q)^s d\omega(p, q) = \int_{J^+} d(p, q)^s d\omega(p, q) \\ &= \underbrace{\int_{E^+} d(p, q)^s d\omega(p, q)}_{=0, \text{ because } d \text{ vanishes on } E^+} + \underbrace{\int_{I^+} d(p, q)^s d\omega(p, q)}_{=0, \text{ because } \omega(I^+) = 0} = 0, \end{aligned}$$

hence a contradiction.

To prove the ‘ $\Leftarrow$ ’ part, suppose there exists  $\omega \in \Pi_c(\mu, \nu)$  with  $\omega(I^+) > 0$ , but nevertheless  $LW_s(\mu, \nu) = 0$ . The latter implies that  $\int_{\mathcal{M}^2} d(p, q)^s d\omega(p, q) = 0$ . But this, in turn, means that

$$\begin{aligned} \int_{I^+} d(p, q)^s d\omega(p, q) &= \int_{J^+} d(p, q)^s d\omega(p, q) - \underbrace{\int_{E^+} d(p, q)^s d\omega(p, q)}_{= 0, \text{ because } d \text{ vanishes on } E^+} \\ &= \int_{\mathcal{M}^2} d(p, q)^s d\omega(p, q) = 0. \end{aligned}$$

But  $d$  is positive on  $I^+$  and so the latter must be an  $\omega$ -null set, which contradicts with the assumption that  $\omega(I^+) > 0$ .

(ii) Let  $\mu_1, \mu_2, \mu_3 \in \mathfrak{P}(\mathcal{M})$  satisfy  $\mu_1 \preceq \mu_2 \preceq \mu_3$ . Let  $\omega_{12}$  and  $\omega_{23}$  be any elements of  $\Pi_c(\mu_1, \mu_2)$  and  $\Pi_c(\mu_2, \mu_3)$ , respectively, and let  $\omega_{123} \in \mathfrak{P}(\mathcal{M}^3)$  be a measure ‘gluing them together’ as specified in the Gluing Lemma. Recall from the discussion following that lemma that  $\omega_{13} := (\text{pr}_{13})_* \omega_{123} \in \Pi_c(\mu_1, \mu_3)$ .

One has the inequality

$$LW_s(\mu_1, \mu_3) \geq \left[ \int_{\mathcal{M}^2} d(p, q)^s d\omega_{12}(p, q) \right]^{1/s} + \left[ \int_{\mathcal{M}^2} d(q, r)^s d\omega_{23}(q, r) \right]^{1/s} \quad (50)$$

which is proven through the following sequence of equalities and inequalities.

$$\begin{aligned} LW_s(\mu_1, \mu_3) &\geq \left[ \int_{\mathcal{M}^2} d(p, r)^s d\omega_{13}(p, r) \right]^{1/s} = \left[ \int_{\mathcal{M}^3} d(p, r)^s d\omega_{123}(p, q, r) \right]^{1/s} \\ &\geq \left[ \int_{\mathcal{M}^3} (d(p, q) + d(q, r))^s d\omega_{123}(p, q, r) \right]^{1/s} \\ &\geq \left[ \int_{\mathcal{M}^3} d(p, q)^s d\omega_{123}(p, q, r) \right]^{1/s} + \left[ \int_{\mathcal{M}^3} d(q, r)^s d\omega_{123}(p, q, r) \right]^{1/s} \\ &= \left[ \int_{\mathcal{M}^2} d(p, q)^s d\omega_{12}(p, q) \right]^{1/s} + \left[ \int_{\mathcal{M}^2} d(q, r)^s d\omega_{23}(q, r) \right]^{1/s}, \end{aligned}$$

where we have used, successively, the definition of  $LW_s$ , the Gluing Lemma (the definition of  $\omega_{13}$ ), the reverse triangle inequality for  $d$ , the reverse Minkowski inequality for integrals [31, Proposition 5.3.1] and, finally, the Gluing Lemma again (the definition of  $\omega_{123}$ ).

By the arbitrariness of  $\omega_{12} \in \Pi_c(\mu_1, \mu_2)$  and  $\omega_{23} \in \Pi_c(\mu_2, \mu_3)$ , inequality (50) immediately yields (49)—one simply has to take the supremum over all  $\omega_{12} \in \Pi_c(\mu_1, \mu_2)$  and all  $\omega_{23} \in \Pi_c(\mu_2, \mu_3)$ .



(iii) By (ii) and the fact that  $\mu \preceq \mu$ , one has  $2LW_s(\mu, \mu) \leq LW_s(\mu, \mu)$ , which is true iff either  $LW_s(\mu, \mu) = 0$  or  $LW_s(\mu, \mu) = +\infty$ .

(iv) To prove ‘ $\Rightarrow$ ’, assume that  $\mathcal{M}$  is chronological. By (i), it suffices to show that for any  $\mu \in \mathfrak{P}(\mathcal{M})$  and for any  $\omega \in \Pi_c(\mu, \mu)$  we must have  $\omega(I^+) = 0$ .

Indeed, proceeding identically as in the beginning of the proof of Theorem 12, we obtain [compare with (46)]

$$\forall f \in \mathcal{L}^1(\mathcal{M}, \mu) \quad \int_{J^+ \setminus \Delta(\mathcal{M})} (f(q) - f(p)) \, d\omega(p, q) = 0. \quad (51)$$

The key now is to use a past volume function  $t^-$  associated with some admissible measure on  $\mathcal{M}$ . Recall that  $t^-$  is causal. Moreover, since  $\mathcal{M}$  is chronological,  $t^-$  is increasing on any future-directed timelike curve (cf. Sect. 2.3). Symbolically:

$$\forall (p, q) \in J^+ \quad t^-(p) \leq t^-(q) \quad \text{and} \quad \forall (p, q) \in I^+ \quad t^-(p) < t^-(q). \quad (52)$$

Substituting  $f := t^-$  in (51) (recall that  $t^-$  is Borel and bounded and hence  $\mu$ -integrable), we can write

$$\int_{E^+ \setminus \Delta(\mathcal{M})} (t^-(q) - t^-(p)) \, d\omega(p, q) + \int_{I^+} (t^-(q) - t^-(p)) \, d\omega(p, q) = 0. \quad (53)$$

By the first property in (52), both integrals in (53) are nonnegative and hence they both must vanish. However, by the second property in (52), the integrand in the rightmost integral is positive on  $I^+$ ; therefore, this integral cannot vanish unless  $\omega(I^+) = 0$ .

The proof of ‘ $\Leftarrow$ ’ is straightforward. Take any  $p \in \mathcal{M}$  and notice that, by assumption,

$$d(p, p) = LW_s(\delta_p, \delta_p) = 0.$$

But this implies (see property (i) of the Lorentzian distance in Sect. 2.3) that  $p \not\ll p$  for any  $p \in \mathcal{M}$ , which means that  $\mathcal{M}$  is chronological.

(v) Suppose that  $LW_s(\mu, \nu) \in (0, +\infty)$  but, nevertheless,  $LW_s(\nu, \mu) > 0$ . By (i), this implies that  $\mu \preceq \nu \preceq \mu$ . By (ii), we can write that

$$0 < LW_s(\mu, \nu) + LW_s(\nu, \mu) \leq LW_s(\mu, \mu).$$

On the other hand, again by (ii), it is also true that

$$LW_s(\mu, \mu) + LW_s(\mu, \nu) \leq LW_s(\mu, \nu),$$

which, since  $LW_s(\mu, \nu)$  is assumed finite, implies that  $LW_s(\mu, \mu) \leq 0$  and we have arrived to a contradiction.  $\square$

Unlike the Lorentzian distance, Lorentz–Wasserstein distances can assume infinite values even in globally hyperbolic spacetimes.

*Example 1.* Consider the  $(1+1)$ -dimensional Minkowski spacetime  $\mathcal{M} := \mathbb{R}^{1,1}$  and fix  $s \in (0, 1]$ . Let  $\mu := \delta_{(0,0)}$  and  $\nu := \sum_{i=1}^{\infty} 2^{-i} \delta_{(2^{i/s}, 0)}$ . Define  $\omega \in \Pi_c(\mu, \nu)$  by<sup>12</sup>

$$\omega := \sum_{i=1}^{\infty} 2^{-i} \delta_{(0,0)} \times \delta_{(2^{i/s}, 0)}$$

and therefore

$$\begin{aligned} LW_s(\mu, \nu) &\geq \left[ \int_{\mathcal{M}^2} d^s d\omega \right]^s = \left[ \sum_{i=1}^{\infty} d\left((0,0), (2^{i/s}, 0)\right)^s 2^{-i} \right]^s \\ &= \left[ \sum_{i=1}^{\infty} \left(2^{i/s}\right)^s 2^{-i} \right]^s = \left[ \sum_{i=1}^{\infty} 1 \right]^s = +\infty. \end{aligned}$$

□

However, Lorentz–Wasserstein distances between two *compactly supported* measures in globally hyperbolic spacetimes are finite.

**Proposition 6.** *Let  $\mathcal{M}$  be a globally hyperbolic spacetime,  $s \in (0, 1]$  and let  $\mu, \nu \in \mathfrak{P}(\mathcal{M})$  be compactly supported. Then,  $LW_s(\mu, \nu) < +\infty$ .*

*Proof.* If  $\Pi_c(\mu, \nu) = \emptyset$ , then trivially  $LW_s(\mu, \nu) = 0 < +\infty$ . Assume then that the set of causal couplings between  $\mu$  and  $\nu$  is nonempty and take any  $\omega \in \Pi_c(\mu, \nu)$ . On the strength of Proposition 4,  $\omega(\text{supp } \mu \times \text{supp } \nu) = 1$ . By assumption, the set  $\text{supp } \mu \times \text{supp } \nu \subseteq \mathcal{M}^2$  is compact. Moreover, by the global hyperbolicity of  $\mathcal{M}$ ,  $d$  is a continuous map and hence it is bounded on that compact set. Therefore,

$$\begin{aligned} \int_{\mathcal{M}^2} d(p, q)^s d\omega(p, q) &= \int_{\text{supp } \mu \times \text{supp } \nu} d(p, q)^s d\omega(p, q) \\ &\leq \max_{\substack{p \in \text{supp } \mu \\ q \in \text{supp } \nu}} d(p, q)^s \int_{\text{supp } \mu \times \text{supp } \nu} d\omega = \left[ \max_{\substack{p \in \text{supp } \mu \\ q \in \text{supp } \nu}} d(p, q) \right]^s \end{aligned}$$

and so, by the arbitrariness of  $\omega$ ,

$$LW_s(\mu, \nu) \leq \max_{\substack{p \in \text{supp } \mu \\ q \in \text{supp } \nu}} d(p, q)^s < +\infty.$$

□

## 6. Outlook

In Definition 2 we proposed a notion of the causal relation between probability measures on a given spacetime  $\mathcal{M}$ . To make sure that the relation is well-defined, we entered the little explored domain on the verge of causality

<sup>12</sup> In fact, it is the only causal coupling between these particular  $\mu$  and  $\nu$ .

and measure theory. The presented formalism can be developed in various directions and applied in both classical and quantum physics.

Firstly, one can try to lower the causality conditions imposed on the spacetime in the theorems presented in Sect. 4. In particular, it would be interesting to see whether the defined relation on  $\mathfrak{P}(\mathcal{M})$  is a partial order for every causal spacetime  $\mathcal{M}$ , or is the assumption (45) in Theorem 12 a necessary one. If the latter holds, one would obtain a new rung of the causal ladder between the causal and distinguishing spacetimes.

A second path of possible development would be to investigate further the notion of a Lorentzian distance in the space of probability measures on a spacetime and the associated topological questions. In Sect. 5 we proposed a notion of the  $s^{\text{th}}$  Lorentz–Wasserstein distance, which is a natural generalisation of the Lorentzian distance between the events on  $\mathcal{M}$ . However, in the optimal transport theory there are other ways to measure distances between probability measures (see for instance [59, p. 97]). It is tempting to see how (if at all) these notions can be adapted to the spacetime framework. This directly relates to the issue of topology on  $\mathfrak{P}(\mathcal{M})$  and its interplay with the semi-Riemannian metric on  $\mathcal{M}$ .

Finally, it is desirable to conciliate between our results and the recent paper of Suhr [55]. In particular, one could check whether Theorems 6, 7, 8, 9 and 10 can be extended to the more abstract, Lorentz–Finsler setting adopted in [55]. This would increase the potential usefulness of our work in the application to the early universe reconstruction problem [14, 28–30].

The first application of the presented theory to the study of causality in quantum theory is discussed in details in [21]. Therein, we focus on the wave packet formalism, which is in common use in atomic, condensed matter [53] and particle physics [12], as an approximation to complicated QFT problems. In this framework, any normalised solution to the Schrödinger equation  $i\hbar\partial_t\psi = \hat{H}\psi$  defines a family of probability measures  $\{\mu_t \in \mathfrak{P}(\mathbb{R}^{n+1})\}_{t \in \mathbb{R}}$  localised on  $t$ -slices, via  $\mu_t = \delta_t \times \|\psi(t, x)\|^2 d^n x$ , where  $d^n x$  is the standard Lebesgue measure on  $\mathbb{R}^n$ . Given such a family of measures on  $(n + 1)$ -dimensional Minkowski spacetime one can, within our formalism, rigorously study the causality of the quantum evolution, i.e. check whether  $\mu_s \preceq \mu_t$  whenever  $s \leq t$  (see also [40]). This allows us to rigorously check the conclusions obtained by Hegerfeldt [35, 36, 39]. We also compare the outcomes with the more recent results on causality in quantum mechanics [3, 5, 9, 60] and seek potential empirical consequences.

To allow for more sophisticated applications in quantum theory one could extend the presented formalism to probability measures with values in some Banach space. The first direct development would be to consider signed measures, which may model both classical and quantum charge densities on a given spacetime. A more challenging task would be to extend the presented causal order to positive-operator-valued measures, what might offer a new insight into the quantum information theory [45].

Finally, let us come back to the original motivation of our preliminary Definition 1. As stressed at the beginning of Sect. 4, it was inspired by the notion of ‘causality in the space of states’ coined in [25]. The partial order

relation considered in [25] is defined on the space of states  $S(\mathcal{A})$  of a  $C^*$ -algebra  $\mathcal{A}$ . If the algebra  $\mathcal{A}$  is commutative, then, by the Gelfand duality, there exists a locally compact Hausdorff topological space  $\mathcal{M}$ , such that  $\mathcal{A} \simeq C_0(\mathcal{M})$ . Then, the Riesz–Markov representation theorem implies that  $S(\mathcal{A}) \simeq \mathfrak{P}(\mathcal{M})$ . Hence, if  $\mathcal{M}$  is a causally simple spacetime, then the two notions of ‘causality for Borel probability measures’ and ‘causality in the space of states’ coincide.

The concept of causality in the space of states was explored [26, 27] in the framework of ‘almost commutative spacetimes’, i.e. for  $C^*$ -algebras of the form  $C_0(\mathcal{M}) \otimes \mathcal{A}_F$ , with  $\mathcal{A}_F$  being a finite dimensional matrix algebra. However, the study therein was limited only to special subclasses of all states, nevertheless yielding interesting results. The theory put forward in the present paper paves the way to unravel the complete causal structure of almost commutative spacetimes. Having in mind that almost commutative spacetimes are utilised to build models in particle physics [57], it is enticing to see whether the extended causal structure imposes any restrictions on quantum probabilities that could be checked experimentally.

## Acknowledgements

ME acknowledges the support of the Marian Smoluchowski Kraków Research Consortium “Matter–Energy–Future” within the programme KNOW and the Foundation for Polish Science within the programme START 2016. TM was partially supported by the Copernicus Center for Interdisciplinary Studies in Kraków and the Foundation ORLEN—DAR SERCA.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- [1] Abdo, A., et al.: Testing Einstein’s special relativity with Fermi’s short hard  $\gamma$ -ray burst GRB090510. *Nature* **462**, 331 (2009)
- [2] Aichmann, H., Nimtz, G.: On the traversal time of barriers. *Found. Phys.* **44**(6), 678–688 (2014)
- [3] Al-Hashimi, M., Wiese, U.-J.: Minimal position-velocity uncertainty wave packets in relativistic and non-relativistic quantum mechanics. *Ann. Phys.* **324**(12), 2599–2621 (2009)
- [4] Amelino-Camelia, G., Ellis, J., Mavromatos, N., Nanopoulos, D., Sarkar, S.: Tests of quantum gravity from observations of  $\gamma$ -ray bursts. *Nature* **393**(6687), 763–765 (1998)
- [5] Barat, N., Kimball, J.: Localization and causality for a free particle. *Phys. Lett. A* **308**(2–3), 110–115 (2003)

- [6] Beckman, D., Gottesman, D., Nielsen, M.A., Preskill, J.: Causal and localizable quantum operations. *Phys. Rev. A* **64**, 052309 (2001)
- [7] Beem, J., Ehrlich, P., Easley, K.: *Global Lorentzian Geometry*. Volume 202 of *Monographs and Textbooks in Pure and Applied Mathematics*. CRC Press, Boca Raton (1996)
- [8] Bernal, A., Sánchez, M.: Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. *Commun. Math. Phys.* **257**(1), 43–50 (2005)
- [9] Berry, M.V.: Causal wave propagation for relativistic massive particles: physical asymptotics in action. *Eur. J. Phys.* **33**(2), 279 (2012)
- [10] Bertrand, J., Puel, M.: The optimal mass transport problem for relativistic costs. *Calc. Var. Partial Differ. Equ.* **46**(1), 353–374 (2013)
- [11] Besnard, F.: A noncommutative view on topology and order. *J. Geom. Phys.* **59**(7), 861–875 (2009)
- [12] Beuthe, M.: Oscillations of neutrinos and mesons in quantum field theory. *Phys. Rep.* **375**(2–3), 105–218 (2003)
- [13] Brenier, Y.: *Optimal Transportation and Applications: Lectures given at the C.I.M.E. Summer School, held in Martina Franca, Italy, September 2–8, 2001. Chapter Extended Monge–Kantorovich Theory*, pp. 91–121. Springer, Berlin (2003)
- [14] Brenier, Y., Frisch, U., Hénon, M., Loeper, G., Matarrese, S., Mohayaee, R., Sobolevskii, A.: Reconstruction of the early universe as a convex optimization problem. *Mon. Not. R. Astron. Soc.* **346**(2), 501–524 (2003)
- [15] Buchholz, D., Fredenhagen, K.: Locality and the structure of particle states. *Commun. Math. Phys.* **84**(1), 1–54 (1982)
- [16] Buchholz, D., Yngvason, J.: There are no causality problems for Fermi’s two-atom system. *Phys. Rev. Lett.* **73**, 613–616 (1994)
- [17] Buscemi, F., Compagno, G.: Non-locality and causal evolution in QFT. *J. Phys. B At. Mol. Opt. Phys.* **39**(15), 695–709 (2006)
- [18] Chruściel, P.T., Grant, J.D.E., Minguzzi, E.: On differentiability of volume time functions. *Ann. Henri Poincaré* **17**(10), 2801–2824 (2016)
- [19] Doplicher, S., Fredenhagen, K., Roberts, J.E.: Spacetime quantization induced by classical gravity. *Phys. Lett. B* **331**(1–2), 39–44 (1994)
- [20] Doplicher, S., Fredenhagen, K., Roberts, J.E.: The quantum structure of space-time at the planck scale and quantum fields. *Commun. Math. Phys.* **172**(1), 187–220 (1995)
- [21] Eckstein, M., Miller, T.: Causal evolution of wave packets. *Phys. Rev. A* **95**, 032106 (2017). doi:[10.1103/PhysRevA.95.032106](https://doi.org/10.1103/PhysRevA.95.032106)
- [22] Einstein, A., Podolsky, B., Rosen, N.: Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**(10), 777–780 (1935)
- [23] Feinberg, G.: Possibility of faster-than-light particles. *Phys. Rev.* **159**, 1089–1105 (1967)
- [24] Foldy, L.L., Wouthuysen, S.A.: On the Dirac theory of spin 1/2 particles and its non-relativistic limit. *Phys. Rev.* **78**(1), 29 (1950)
- [25] Franco, N., Eckstein, M.: An algebraic formulation of causality for noncommutative geometry. *Class. Quantum Gravity* **30**(13), 135007 (2013)

- [26] Franco, N., Eckstein, M.: Exploring the causal structures of almost commutative geometries. *Symmetry Integr. Geom. Methods Appl.* **10**, 010 (2014). (Special Issue on Noncommutative Geometry and Quantum Groups in honor of Marc A. Rieffel)
- [27] Franco, N., Eckstein, M.: Causality in noncommutative two-sheeted space-times. *J. Geom. Phys.* **96**, 42–58 (2015)
- [28] Frisch, U., Matarrese, S., Mohayaee, R., Sobolevski, A.: A reconstruction of the initial conditions of the universe by optimal mass transportation. *Nature* **417**(6886), 260–262 (2002)
- [29] Frisch, U., Podvigina, O., Villone, B., Zheligovsky, V.: Optimal transport by omni-potential flow and cosmological reconstruction. *J. Math. Phys.* **53**(3), 033703 (2012)
- [30] Frisch, U., Sobolevskii, A.: Application of optimal transport theory to reconstruction of the early universe. *J. Math. Sci.* **133**(4), 1539–1542 (2006)
- [31] Garling, D.J.H.: *Inequalities: A Journey into Linear Analysis*. Cambridge University Press, Cambridge (2007)
- [32] Gell-Mann, M., Goldberger, M.L., Thirring, W.E.: Use of causality conditions in quantum theory. *Phys. Rev.* **95**, 1612–1627 (1954)
- [33] Haag, R.: *Local Quantum Physics: Fields, Particles, Algebras*. Theoretical and Mathematical Physics. Springer, Berlin (1996)
- [34] Hawking, S.W.: Chronology protection conjecture. *Phys. Rev. D* **46**, 603–611 (1992)
- [35] Hegerfeldt, G.C.: Remark on causality and particle localization. *Phys. Rev. D* **10**, 3320–3321 (1974)
- [36] Hegerfeldt, G.C.: Violation of causality in relativistic quantum theory? *Phys. Rev. Lett.* **54**, 2395–2398 (1985)
- [37] Hegerfeldt, G.C.: Causality, particle localization and positivity of the energy. In: Bohm, A., Doebner, H.-D., Kielanowski, P. (eds.) *Irreversibility and Causality Semigroups and Rigged Hilbert Spaces*, Volume 504 of *Lecture Notes in Physics*, pp. 238–245. Springer, Berlin (1998)
- [38] Hegerfeldt, G.C.: Particle localization and the notion of Einstein causality. In: Horzela, A., Kapuścik, E. (eds.) *Extensions of Quantum Theory*, pp. 9–16. Apeiron, Montreal (2001)
- [39] Hegerfeldt, G.C., Ruijsenaars, S.N.M.: Remarks on causality, localization, and spreading of wave packets. *Phys. Rev. D* **22**, 377–384 (1980)
- [40] Miller, T.: Polish spaces of causal curves. *J. Geom. Phys.* **116**, 295–315 (2017). doi:[10.1016/j.geomphys.2017.02.006](https://doi.org/10.1016/j.geomphys.2017.02.006)
- [41] Miller, T.: On the causality and  $K$ -causality between measures. Preprint [arXiv:1702.00702](https://arxiv.org/abs/1702.00702) [math-ph] (2017)
- [42] Minguzzi, E.: Time functions as utilities. *Commun. Math. Phys.* **298**(3), 855–868 (2010)
- [43] Minguzzi, E., Sánchez, M.: The causal hierarchy of spacetimes. In: Alekseevsky, D.V., Baum, H. (eds.) *Recent Developments in Pseudo-Riemannian Geometry*. ESI Lectures in Mathematics and Physics, pp. 299–358. European Mathematical Society, Helsinki (2008)
- [44] Moretti, V.: Aspects of noncommutative Lorentzian geometry for globally hyperbolic spacetimes. *Rev. Math. Phys.* **15**(10), 1171–1217 (2003)

- [45] Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2010)
- [46] O'Neill, B.: Semi-Riemannian Geometry with Applications to Relativity. Academic Press, New York (1983)
- [47] Pawłowski, M., Paterek, T., Kaszlikowski, D., Scarani, V., Winter, A., Żukowski, M.: Information causality as a physical principle. *Nature* **461**(7267), 1101–1104 (2009)
- [48] Penrose, R.: Techniques of Differential Topology in Relativity. Volume 7 of CBMS–NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia (1972)
- [49] Peres, A., Terno, D.R.: Quantum information and relativity theory. *Rev. Mod. Phys.* **76**, 93–123 (2004)
- [50] Ringström, H.: The Cauchy Problem in General Relativity. ESI Lectures in Mathematics and Physics. European Mathematical Society, Helsinki (2009)
- [51] Rudin, W.: Real and Complex Analysis, 3rd edn. McGraw-Hill, New York (1987)
- [52] Srivastava, S.M.: A Course on Borel Sets, Volume 180 of Graduate Texts in Mathematics. Springer, New York (2008)
- [53] Strange, P.: Relativistic Quantum Mechanics: with Applications in Condensed Matter and Atomic Physics. Cambridge University Press, Cambridge (1998)
- [54] Streater, R.F., Wightman, A.S.: PCT, Spin and Statistics, and All That. Princeton Landmarks in Mathematics and Physics. Princeton University Press, Princeton (2000)
- [55] Suhr, S.: Theory of optimal transport for Lorentzian cost functions. Preprint [arXiv:1601.04532](https://arxiv.org/abs/1601.04532) [math-ph] (2016)
- [56] Thaller, B.: The Dirac Equation, Volume 31 of Theoretical and Mathematical Physics. Springer, Berlin (1992)
- [57] van Suijlekom, W.D.: Noncommutative Geometry and Particle Physics. Mathematical Physics Studies. Springer, New York (2015)
- [58] Villani, C.: Topics in Optimal Transportation. Graduate Studies in Mathematics. American Mathematical Society, Providence (2003)
- [59] Villani, C.: Optimal Transport: Old and New. Volume 338 of Grundlehren der mathematischen Wissenschaften. Springer, Berlin (2008)
- [60] Wagner, R., Shields, B., Ware, M., Su, Q., Grobe, R.: Causality and relativistic localization in one-dimensional Hamiltonians. *Phys. Rev. A* **83**, 062106 (2011)
- [61] Willard, S.: General Topology. Addison-Wesley, Reading (1970)
- [62] Winful, H.G.: Tunneling time, the Hartman effect, and superluminality: a proposed resolution of an old paradox. *Phys. Rep.* **436**(1–2), 1–69 (2006)

Michał Eckstein

Faculty of Physics, Astronomy and Applied Computer Science

Jagiellonian University

ul. prof. Stanisława Łojasiewicza 11

30-348 Kraków

Poland

e-mail: [michal.eckstein@uj.edu.pl](mailto:michal.eckstein@uj.edu.pl)

Michał Eckstein and Tomasz Miller  
Copernicus Center for Interdisciplinary Studies  
ul. Szczepańska 1/5  
31-011 Kraków  
Poland

Tomasz Miller  
Faculty of Mathematics and Information Science  
Warsaw University of Technology  
ul. Koszykowa 75  
00-662 Warsaw  
Poland  
e-mail: [T.Miller@mini.pw.edu.pl](mailto:T.Miller@mini.pw.edu.pl)

Communicated by James A. Isenberg.

Received: April 18, 2016.

Accepted: January 31, 2017.